

Does Britain want a free press?

The British government's endless disputes with its media can be resolved only by a reform of policies on official information. More regulation leads the wrong way.

BRITAIN, fond of describing itself as the oldest of democracies (earlier Scandinavian precedents notwithstanding), is in danger of lapsing into undemocratic ways in its treatment of its press. So much is evident from the way in which, in Britain, government relations with newspapers, broadcasting organizations and publishers have themselves become public issues. The latest development is last week's row between the government and the Independent Broadcasting Authority (which regulates commercial television stations) and the self-regulating British Broadcasting Corporation (the independent broadcasting organization financed by a levy on people who own television receivers) about the shooting of three alleged terrorists in Gibraltar in March. Both organizations are in trouble with the government for having broadcast films dealing with the circumstances of the shootings. Mrs Margaret Thatcher, the prime minister, called the first broadcast "trial by television". Sir Geoffrey Howe, the foreign secretary, much respected for level-headedness, has spoken of his "deep regret" that the BBC went ahead with its broadcast despite a request from the home secretary, Mr Douglas Hurd, that the film should be suppressed. But this dispute, which threatens further to corrode relations between the government and the press, stems largely from the British government's opinion of the uses to be made of official information which cloud its whole relationship with the press.

It seems to be common ground that the three people shot in Gibraltar were members of the Irish Republican Army (IRA) who were planning to set off a bomb and that they were shot and killed by members of a British Army unit called the Special Air Services wearing civilian clothes. Further information of the circumstances may come to light at the inquest on the deaths to be held in Gibraltar at some date after 20 June. Even though the IRA, with customary ghoulish bombast, has described its dead members as having been "on active service", there is legitimate public concern that they were killed on public streets when arrest might have been more appropriate. That is why both newspapers and the television broadcasters have been combing Gibraltar for evidence of what may have happened.

Protests

The government's protests, directed only at the television organizations, turn on the complaint (which may be valid) that interviewing supposed witnesses of the shootings will, by rehearsal, contaminate the evidence they may give to the intended inquest. One snag is that Gibraltar, while technically a British possession, has a separate legal system and that the British law of contempt that would prevent journalists from anticipating the outcome of a British legal process do not, on this occasion, apply. Another is that the government itself, in the immediate aftermath of the shootings, felt free to say that the three IRA members were adequately warned, and were shot only when they moved suspiciously. If the shootings had taken place in Britain, government ministers and spokesmen would no doubt have been more circumspect, for fear of prejudicing the jury at the inevitable inquest.

The effect on the British population of government statements, if uncontaminated by contradictory reports, would thus

have been to promulgate one view of the events in Gibraltar which, given that an inquest is not a trial (there are no defendants) and that the evidence presented to an inquest's jury is determined by the coroner in charge (as well as by the resources at his disposal), might never have been challenged. Correct though the opinion may be that rehearsing witnesses in advance runs the risk of contaminating their later evidence, how is that to be measured against the danger that a legitimate public issue (the proper rules of engagement of the security forces in confrontation with suspected terrorists) will now be buried? The question is all the more insistent because of the collapse, two years ago, of the inquiry into allegations that security forces in Ulster in the early 1980s were allowed to shoot suspected terrorists without warning.

Embarrassed

All governments have a natural interest in presenting their own doings as favourably as possible, and are also often embarrassed by the covert doings of their security and intelligence services. (The issue in this case is not whether British soldiers had a right to operate in Gibraltar against the IRA, but whether they were trigger-happy.) Suspicions that the British government's pressure on the television organizations stems from more than mere concern for the propriety of next month's inquest must be reinforced by its three-year prosecution in the British and Australian courts of the case against a book by a retired intelligence officer, with the result that sensational (but possibly incorrect) allegations of dirty tricks by British intelligence still cannot be openly reported and publicly discussed.

Sadly, for democratic governments, there is no escape from the dilemma caused by the continued existence of a free press — the risk (indeed, the certainty) that well-intentioned policies will be criticized in ways that appear disloyal (and which may also be mistaken). British governments (not just the present government) nevertheless make needlessly heavy weather of the problem by their assiduous management of official information, most of which is secret but much of which is leaked selectively, in a manner that cannot be verified and which may even be intended to mislead. (It is not so long ago that a spokesman at the Ministry of Defence deliberately misled this journal with incorrect information about the innocuous matter of the publication date of the Swinnerton-Dyer report on last year's great storm.) Can it be any wonder that many British newspapers and other organs of information are able to purvey surmise and fairy-stories as if they were fact?

The irony is that, having created a climate of misinformation, the government is impatient with the irresponsibility its policies have engendered. Inevitably, the House of Commons bristles with drafts of private members' bills intended to "curb" the press, while the voluntary Press Council dealing with public complaints is hard-pressed to find a new chairman likely to seem tough enough to satisfy the common public opinion that "something must be done" and the opinion within the newspaper industry that, otherwise, it will have legislation to contend with instead.

Yet the remedy lies not with the press and the other organs of



communication. Ten years have passed since a previous government promised to reform the Official Secrets Act, and seven since the present government failed to win support for what would have been a repressive amendment. Now, a new statement of policy is promised for June or July. It is to be hoped that something will materialize, and that it will betoken a commitment by the British government to the principle that official information should ordinarily be open information even when its appearance is uncomfortable. But that will not remove the difficulty about Gibraltar, where the neat (but, for the government, unwelcome) way of securing silence in advance of a public hearing is to arrange for a judicial inquiry to be held in Britain.

Give charlatans a break

Reports of astrologers at the White House will, as usual, damage the reporters.

PRESIDENT Ronald Reagan has responded with accustomed calm and dignity to the charge by Mr Donald T. Regan, published in book form earlier this week, that many of the decisions made during the presidency now moving towards its close have been influenced by astrology. There could be no more telling proof of Regan's unsuitability for the post he occupied as chief of staff at the White House than that, on some pages of his book, he should tell of President Reagan's universal kindness and, on others, fail to recognize a kindly man's unwillingness to offend even his chief of staff, when rejecting his advice, by taking refuge in what the stars foretell.

Readers of this disgruntled adviser's book will readily see that the President, when advised that he should go to Pittsburgh (for example) next Tuesday (say), would not have made the natural and cruel reply "That's a dumb idea, Don", but would instead have vaguely referred to the unfavourable conjunction of the stars and planets on the suggested day. For Regan to have failed to see this explanation without prompting is proof that he cannot have understood the job required of him at the White House.

Indeed, if Regan's account of the consultations with astrologers is to be believed, it will most probably be understood for what it is — a telling demonstration by the imaginative and radical president, who launched the federal budget deficit and the Strategic Defense Initiative, that administrative advice on what to do when should habitually be taken with a large pinch of salt. The advisers, after all, have usually no better reason for sending the president to Pittsburgh next Tuesday than the astrologers have for sending him there (or somewhere else) another day. What better way of putting the advisers in their place than by arranging for an external source of random advice on the tedious details of the presidential life? What more galling for the professional soothsayers who function as chiefs of staff than that the alternative advice should come from people liberated from the notion that questions such as when to go to Pittsburgh require rational decision?

In any case, as Regan must acknowledge, the alternative advice has worked. Nothing has gone disastrously wrong during the past eight years. Occasional episodes of acute discomfort, such as the discovery of the Irangate scandal two years ago, have not proved to be the cataclysms they might have been. Even the emergence of the United States as the most heavily indebted debtor nation has not (yet) interrupted the growth of the second-richest economy in the world. For all that mean-minded Regan knows, there could have been much more serious trouble during the past few years if the astrologers had been ignored. One of the perennial obstacles astrologers face is that the charge of charlatanry cannot be shaken off so long as they issue advice which misguides people then follow. A more charitable observer than Donald Regan, more willing to follow the example of his betters, might powerfully have helped to rehabilitate this much-maligned community whose only offence is to have managed to make mystique saleable. □

Selling nuclear power

A new government committee has drawn attention to problems in selling nuclear power stations.

SINCE the publication in February of the British government's plan for the sale of the electricity supply industry, little has been heard of the nuclear power component of that enterprise except that the Nuclear Installations Inspectorate (with statutory power to regulate the safety of nuclear plants) has been issuing gloomy forecasts of the predictable end of the civil nuclear power station first planned more than 30 years ago. But now the Advisory Council on Science and Technology (ACOST) has raised the issue of what will happen after the electricity industry has been privatized (see p. 108).

ACOST itself is a new creation, less than a year old, which has plainly judged it prudent to get some kind of study completed quickly. It has embarked on and finished an examination of the future of the British nuclear power industry now that it has been decided that a proposed pressurized-water reactor station should be built at Sizewell, on the coast of Suffolk.

So far as it goes, the report of the study is down-to-earth and workmanlike. Having at last settled on a reactor design, ACOST says, the British nuclear industry had better stick with it "for many years". Having agreed that the new plant should be a foreign design (by Westinghouse of the United States), there may be some anxiety about the degree to which British funds are used to buy components and services from elsewhere, but this will be serious only if the European Communities succeed in removing "some of the energy sector trade restrictions" now extant. But there will be little effect on employment, while there is unlikely to be any obvious shortage of skilled people. In the aftermath of February's declaration that the electricity industry is to be sold, the report is necessarily somewhat archaic.

This is bad luck for which ACOST cannot be blamed. But there remain several important issues about nuclear power to be decided before the supply industry is eventually sold off. For, hitherto, British nuclear power has been a public enterprise, supported on the research and development side directly from taxes and otherwise by revenues collected from electricity consumers. There is nothing in the government's sketchy account of how privatization will be accomplished to suggest what its continuing role in the support of nuclear power will be, although it seems inevitable that it must continue to support work in radiological protection, waste disposal and safety improvements.

Both ACOST and the rumour-mills support the notion that fast-reactor development will be offered to the electricity utilities and, if not embraced, abandoned. The existing Central Electricity Generating Board (CEGB) has been saying that it will be generating 40 per cent of its electricity from nuclear reactors 30 years from now, but its commercial successors may sing a different tune. The long-term economic advantage of nuclear power relative to coal is that the smaller proportion of its total cost attributable to labour insulates it against the rising cost of sharing increased national prosperity, but this may not be evident to new shareholders. That is why there is a case for using a small part of the proceeds of the impending privatization to secure a tapering programme of research.

The management of nuclear projects is a different kettle of fish. ACOST has a judicious discussion of the several unsatisfactory attempts over the past two decades to build up an independent organization capable of managing nuclear contracts. The National Nuclear Corporation, the present organization, has the status of a public company which is neither capable of undertaking to build a nuclear power station for a customer such as CEGB nor as acting as an agent for the customer. ACOST says this is unsatisfactory (which is true), but that privatization must see the creation of a strong design and procurement organization. But that, surely, is for the newly privatized organizations to decide. □

French AIDS research effort comes under attack

■ European AIDS federation launched

■ French AIDS finances attacked

Paris

PROFESSOR Luc Montagnier last week announced in Paris the launch of a European Federation for Research on AIDS. But, on a national level, he feels that the French government's policy on AIDS research is "inconsistent" and confesses to being "very concerned" about the financial future of his own laboratory at the Pasteur Institute.

The European Federation for Research on AIDS (FERS), which has so far received support from the West German Korber Prize and the French Mutuelle civil service pension fund, aims to stimulate European cooperation between scientists actively engaged on AIDS research. Montagnier feels this will minimize conflict with other international organizations — such as the French American AIDS Foundation and the World AIDS Foundation (see *Nature* 330, 507; 1987) — which also finance information campaigns and paramedical workers, especially in Africa.

The federation whose membership includes representatives from Britain, West Germany, Italy and Switzerland, as well as France, has set itself four major research themes: virology (for example to study the role of regulatory proteins in the HIV-1 molecule), DNA (including research on HIV-2), immunology (mechanisms of infection) and clinical epidemiology.

The federation hopes to attract financial support from major pharmaceutical companies which, says Montagnier, have not been pulling their weight.

If the spirit of international cooperation and coordination within AIDS research is encouraging, Montagnier is less happy about levels of state funding this year. "Last year the government provided FF100 million (£10 million) for AIDS research. This year we have received only FF20 million. It's ridiculous." Although a reduction in the national budget was expected, because the first award included FF40 million for non-recurrent construction and equipment costs, Montagnier does not know if, or how much, extra money will be forthcoming. "We can't live off promises. What is needed is a 3–5 year plan, not the squeeze-box effect of massive funding in election years, followed by cuts when the government changes."

Montagnier says his own laboratory is running out of money and is currently dependent solely on (reduced) government grants. The Pasteur Institute

receives money from the French American AIDS Foundation and from fund-raising drives, but Montagnier says that little of this money comes to his laboratory, contrary to what may be supposed.

Professor Jean-Paul Lévy, who coordinates French AIDS research within the public sector (see *Nature* 328, 191; 1987) is more optimistic. "The FF20 million from



Montagnier — French AIDS policy critic.

the government is a first instalment for 1988, although it is true we do not know how much more will be provided. We expect to get a bit less this year, but, naturally, if the total fell below FF60 or FF70 million the national programme would become meaningless. But it should not be forgotten that salaries are already covered by the government's overall research budget."

A real problem, on a national scale, says Lévy, is the cutback in research appointments (see *Nature* 329, 380; 1987) a view echoed by Montagnier as well as by unions representing researchers. While there is a slight preference for the major research organizations, such as CNRS and INSERM, to take on postdoctoral candidates in life sciences who want to work on AIDS, says Montagnier, "competition is great and other areas of research inevitably suffer". Money is available, however, to invite foreign researchers, especially within virology, an area that has been relatively neglected in France. But, says Lévy, there has been little response from young European virologists so far — an appeal quite specifically aimed at Britain.

Peter Coles

No visas for AAAS

Oxford

South African Minister of Home Affairs Stoffel Botha announced in Cape Town on 26 April that he had refused visas to a 7-member delegation of the American Association for the Advancement of Science. The delegation had intended to visit South Africa to investigate how the emergency legislation had affected the provision of medical services in the country. M.C.

Plastic wrapping ban

Washington

SUFFOLK County, Long Island has become the first US county to attempt to ban plastic wrapping materials and replace them with biodegradable paper materials. From July 1989, the use of containers and bags made of polystyrene and polyvinylchloride will no longer be permitted in the county's grocery and food establishments. The county was one of the first to bar non-returnable bottles. A.A.

New OUP chief named

London

Professor Sir Roger Elliott, head of the theoretical physics department at the University of Oxford, has been appointed secretary to the delegates of Oxford University Press and chief executive for five years from 1 September 1988. M.S.

Student education tax

Sydney

AUSTRALIA'S Wran Committee has come out in favour of a tax surcharge for those receiving higher education. The committee, under Sir Neville Wran, has been trying to find ways of raising extra funds to pay for a planned increase in student enrolment (see *Nature* 332, 773; 1988).

Students will be expected to repay 20 per cent of the costs of their higher education through a tax surcharge, thought to be less than 5 per cent. The tax will apply to students who do not complete their education as well as to those who graduate. In medicine, dentistry, veterinary science and law, annual education costs amount to around A\$53,000. T.E.

Rabies in Finland

London

RABIES, unreported in Finland since 1959, has now broken out again, having reached the country from the Soviet Union. A massive inoculation programme has been launched, including, according to Agriculture Minister Toivo T. Pohjala, the oral vaccination of wild animals such as raccoon dogs. Seven new research posts to deal with rabies have been established by the Agriculture Ministry. Although only a few cases of the disease have so far been reported, there is concern that two — a fox and a dog — occurred well outside the border danger zone. V.R.

"Shot in the arm" for Japan's program

Tokyo

JAPAN'S Human Frontiers Science Program received a shot in the arm on 29 April when scientists from the seven Western summit nations (Japan, Canada, Britain, France, Italy, West Germany and the United States) and the European Community met in Bonn and declared enthusiastic support for the project and called for its implementation at the earliest possible date. But it will be the end of this year before it becomes clear how much Japan is prepared to put into the programme.

The Bonn "Wise Men's Conference", supported by the Japan Science Foundation and the Alexander von Humboldt Foundation, follows a meeting of the summit nation scientists in Tokyo in March at which it was decided that Frontiers should promote international basic research on the brain and the molecular mechanisms of biological functions by awarding research grants and fellowships through international peer review and by staging frequent workshops. The scientists estimated that an annual budget of about \$80–\$100 million would be appropriate in the initial phase.

Japan's Prime Minister, Noboru Takeshita, is expected to raise Frontiers at the next summit in Toronto in June. Takeshita has expressed support for the programme and it is in line with his recent calls for increased trilateral collaboration between Japan, Europe and the United States. But it is unlikely that he will be able to put any money on the table at Toronto.

Although Frontiers was initiated by the Ministry of International Trade and Industry (MITI), the chief responsibility for negotiating the budget has swung to the Science and Technology Agency (STA). MITI is taking a backstage role in the project partly because Western suspicions about a trade and industry ministry being involved in a basic research project was giving the programme a poor image. But the agency does not have the same clout as MITI. Furthermore, the Ministry of Finance, which will allot the budget for Frontiers, is divided into sections for each ministry, and while MITI's section in the Finance Ministry is thoroughly familiar with the programme after two years of negotiations, there are some newcomers in STA's section and budget negotiations are expected to be prolonged. STA and MITI will submit their budget requests to the Finance Ministry at the end of August.

In these circumstances, it seems unlikely that other summit nations will make any financial commitment at Toronto. But MITI officials are confident that in the long term at least France and West Germany will contribute funds to the programme.

David Swinbanks

Little hope of Yuppiedom for benchbound scientists

London

THE principle discovery of the alchemists, that you won't get rich at the laboratory bench, still holds true — at least for British scientists. A study by an independent group of salary analysts, Income Data Services (IDS)*, confirms what scientists here have known for years: "The pay levels for professional engineers and scientists . . . compare unfavourably with the earnings of other professional groups."

According to the study, the average gross annual salary for full-time males in science, engineering and technology was, at April last year, £13,884 (\$25,825). This compares with £17,056 for professionals supporting management and administration; £18,876 for personnel officers and managers; £19,188 for marketing and sales managers and executives; and £24,700 for finance, insurance and tax specialists. For women scientists the situation is even more dire, with the average annual gross salary at only £9,261. The report offers no explanation for this.

For those scientists who resist the poaching efforts of the financial sector, a move into management and administration usually means a jump of several rungs up the salary ladder. IDS found last year that for most research and development specialists without management responsibility there was a salary bar of between £16,000 and £20,000. Earlier this year, a survey by the Royal Society of Chemistry showed that median earnings for its members in general management were £27,800; research and development pri-

marily as a manager or administrator, £22,000; marketing and sales, £21,000; and research and development without substantial management or administrative responsibilities, £15,300. Similarly, the Institution of Mechanical Engineers found that general administration paid well with earnings of around £27,200 for the 50–54 age group. The Institution of Civil Engineers recorded £20,350 for members in marketing and sales, but only £15,000 for those in development and design.

The IDS study found that in general the public sector pays less well than the private sector, and thus has had to introduce several counter measures to poaching by private firms. In particular, new salary structures have been introduced to allow merit payments and opportunities for faster promotion. Poaching is particularly acute in the financial sector. "There are signs that newly qualified scientists and engineers, with their numerate, logical and analytical minds, are being snapped up by employers who value them for these qualities rather than for their specific technical training", the IDS says.

Relatively high median salaries for university professors and lecturers (£21,000 for chemists) "merely reflect the fact that the age distribution is so different from that in other occupations, with the university employees concentrated in the older age groups," the report says.

Simon Hadlington

*IDS study 408, *Engineers' and Scientists' Pay* (193 St John Street, London EC1V 4LS)

Euroregulations proposed for modified organisms

London

No release of a genetically modified organism will be allowed within any country of the European Communities (EC) without the other members having a chance to object, if new proposals by the Commission of the EC are adopted. The Commission itself will be the final court of authority should any country object on grounds of safety to the sale on its territory of an organism that has received national approval elsewhere.

The Commission's proposals are an attempt to 'harmonize' regulations within Europe. They deal not only with the release of genetically modified organisms but also with their use in contained conditions, ranging from academic laboratories to production plants.

Each use of genetically modified organisms will have to be notified to a relevant

national authority, except in the case of the contained use in non-industrial laboratories or microorganisms classified as 'minimal hazard'. For some uses, notification will be largely a matter of routine, but as the potential risk attached to the experiment or process increases so will the degree of endorsement necessary.

The Commission's proposals will now be subject to scrutiny by the Council of Ministers and the European Parliament, and will be scrutinized by environmental and industrial groups. In view of the widely divergent current position of EC countries on regulations, ranging from the highly permissive (such as Italy) to the highly restrictive (such as may be adopted in West Germany — see *Nature* 332, 672; 1988), it is unlikely that the proposals will be converted into a binding directive in the immediate future. **Peter Newmark**

Berkeley computer sleuth outwits West German hacker

Munich

RESEARCHERS at Lawrence Berkeley Laboratories (LBL) in California tracked an intruder in their computer system for almost a year as he sifted through hundreds of US systems seeking classified data. The intruder was finally located in West Germany, but prosecution has foundered for lack of evidence, although the FBI is still looking into the possibility of espionage.

The incident has revealed weak points in computer operating systems and security procedures as well as in West German law on computer espionage. But by treating their investigation as a "valuable research project", Berkeley astronomer Clifford Stoll and his colleagues have gained experience of tracing 'hackers' which may help prevent future break-ins.



Stoll—investigating computer espionage.

Stoll first detected the intruder in August 1986 when a discrepancy arose in the LBL computer's internal accounting records. Soon afterwards, the National Computer Security Center (NCSC) at Fort Mead, Maryland, notified LBL that someone using an LBL account had tried to access its computer, which turned out to be the same person.

LBL is thought to have been the target because of its proximity, in name and location, to Lawrence Livermore Laboratories, where classified defence projects are carried out. LBL does no classified research. Stoll, at first believing the intruder to be a Berkeley student, spent ten months tracing the hacker to Hannover, West Germany, in June last year.

The hacker had tried to gain access to about 450 computer systems in the United States, Japan and West Germany, most of them used for military work. Operating systems breached included AT&T's UNIX, Digital Equipment's VMS and International Business Machines' VM-TSO. In total, the intruder breached more than 30 of the systems, but Stoll says

that LBL researchers monitored "every keystroke", notifying the violated systems of the intrusions.

The eventual give-away was the communications delay. Working with US and West German authorities, Stoll laid a trap using made-up data about the US Strategic Defense Initiative (SDI). The intruder took the bait in June 1987, staying connected with the LBL computer for over an hour—long enough for the West German authorities to trace and arrest him.

An account of Stoll's project has been published in the May issue of *Communications of the Association for Computer Machinery*, a respected computer science journal, which blames sloppy security procedures for many of the break-ins.

But the case against the alleged hacker has had to be dropped because he was not caught in the act. A local court found that the telephone trace back to his apartment was insufficient evidence to prove he had

been the culprit. West German authorities returned the suspect's confiscated computer and memory diskettes unread because they never received the court's permission to check for the LBL files.

Bremen public prosecutor Hans-Georg von Bock und Polach says that officials' hands are tied in part because merely entering a computer system is not illegal in West Germany. Authorities could have gained access to data coming over the telephone with a wiretap, but such a procedure would only have been approved if national security—of either West Germany or the United States—had been thought to be at risk.

The continuing FBI investigation has been triggered by a letter to Stoll from Pittsburgh enquiring about some of the material in the trumped-up SDI file; the author identified himself as an "international arms dealer".

Stoll says that the implications of the case for open international computer networks are dire, remarking that funding agencies have already eliminated some international links because of break-ins.

Steven Dickman

Top French mathematician turns down lucrative Crafoord prize

Paris

FRENCH mathematician Alexander Grothendieck, whose revolutionary work in algebraic geometry won him the Fields medal in 1966, has turned down the Crafoord prize he was to share with his former student, Pierre Deligne. In his letter to the Swedish Royal Academy of Sciences published in the French newspaper *Le Monde* on 4 May, Grothendieck not only questions the justification for lucrative honours but also attacks the morality of contemporary mathematics.

Grothendieck, who is due to retire in October from his post at the University of Montpellier, is regarded highly by his colleagues. But, according to Deligne—who has accepted the award—the snub comes as no surprise. "He is a very obstinate man. When he was awarded the Fields medal, he did not go to Moscow to receive it."

Since 1970, Grothendieck has withdrawn from active research. His outspoken views on world ecology have subjected him to ridicule, especially after a stormy period at the Collège de France when his non-mathematical preoccupations disappointed those who had nominated him. This withdrawal has apparently led to the disillusion principally behind his decision not to accept the Crafoord prize.

"In the past two decades", he says "the ethics of the scientific milieu (at least as far as mathematics is concerned) have decayed to the point where pillage, pure and simple, between colleagues (especi-

ally at the cost of those least able to defend themselves) has become almost the general rule. In any case it is tolerated by all, even in the most flagrant and iniquitous cases."

He also feels that such awards serve no useful purpose: "fertility is rewarded by progeny, not by honours". Even the financial reward—the Crafoord prize is worth \$270,000—is unnecessary, he says. Such prizes are usually awarded to high-level researchers who already have "material well-being and scientific prestige".

The publication of Grothendieck's letter has taken the French mathematics community by surprise, even if his views are well known. Fearful that the frank rebuttal will displease the Swedish Royal Academy, leading mathematicians are seeking a representative to apologise.

But Pierre Deligne feels the attack will have little real impact on international mathematics. "Of course plagiarism exists within mathematics, but probably less so than in other fields of science. Sometimes, when a result has been found, the author feels no need to publish". Another former colleague, Jean-Pierre Ferrier, shares this view. "Grothendieck is a man of extraordinary generosity and has left a permanent mark both on mathematics and his students. But his view of scientific research has always been idealistic. The reality of competitiveness was always there, but perhaps it is only now that he sees it".

Peter Coles

Defence mounted for research in biological warfare

Washington

MILITARY and government officials squared off against environmentalists and medical professionals over the propriety of the Army's biological defence programme at a congressional hearing last week. At issue is whether the Army has crossed the fine line between defensive and offensive weapons research.

Controversy over the Army's biological defence programme has come to a head since the Army lost a lawsuit in 1985 brought by Jeremy Rifkin's Foundation on Economic Trends. As a result, the Army had to file an environmental impact assessment of its plans for a high-containment biological research facility at the Dugway Proving Ground, Utah. Plans for the facility include the testing of protec-

tive gear against pathogen aerosols and the development of vaccines and treatments to protect troops against biological weapons on the battlefield. The Army filed its assessment with the Environmental Protection Agency three months ago.

The agency has responded by requesting the Army to provide more information on the use of non-pathogenic 'simulant' organisms as options to reduce risk; to detail more fully security measures to safeguard the installation against sabotage or terrorism; and to provide additional justification for the highest biosafety level status, if it contends that research requiring such maximum containment will not be conducted. Rifkin has also pressured the Army into preparing an environmental impact statement, due this month, covering all the Army contractors conducting biological weapons research.

The United States cannot produce or stockpile biological weapons under the 1972 Biological Weapons Convention, signed by 103 nations including the Soviet Union. But because the treaty contains no provisions for verification, it allows research on virulent microorganisms and toxins for "prophylactic, protective, or other peaceful purposes".

The United States has accused the Soviet Union of using mycotoxins, or

'yellow rain', in Afghanistan and South-East Asia, and points to a mysterious outbreak of anthrax in Sverdlovsk in 1979 as evidence that the Soviets are developing biological weapons. The Soviets have repeatedly denied the allegations.

In his testimony before the joint meeting of three congressional subcommittees, Thomas J. Welch, deputy assistant to the Secretary of Defense, said the number of countries capable of waging war using biological weapons has increased to 10 in recent years. Because several of them are unfriendly to the United States and are on the State Department's list of countries that harbour international terrorists, he said "we . . . have the responsibility to the soldier" to develop defence mechanisms.

But public health witnesses, including Anthony Robbins of the Committee for Responsible Genetics and Jay A. Jacobson of the University of Utah School of Medicine, argued that the diversity of agents that could be used as biological weapons and the many methods for deploying them prevent the realistic development of defence strategies. Instead, they countered, the Army's work in testing actual pathogens against potential defences could be construed by other countries as offensive research. They also argued that the \$300 million project planned for Dugway is wasteful, considering that no effective defence can be developed, and that the risks, should one of the organisms being studied escape, are too high.

Carol Ezzell

South African minister backs down

Oxford

SOUTH AFRICA'S Minister of National Education, F.W. de Klerk, seems finally to have backed down in his efforts to impose subsidy conditions on the country's universities. This follows rulings of both the Cape and Natal divisions of the Supreme Court that the regulations imposing subsidy conditions, promulgated in October last year, were *ultra vires*.

It was feared that de Klerk might respond by amending the Universities' Act, but last week, in his budget vote in parliament, he said that the letters sent to the universities last October outlining the subsidy conditions were to be withdrawn. Instead he announced that the government would enter into negotiations with the universities to achieve consensus on the issue of campus discipline. The universities of Cape Town and the Witwatersrand have welcomed his proposals.

There has been speculation that the real reason de Klerk has backed down may relate to the fact that the different universities fall under the control of several ministers of education (see *Nature* 331, 104; 1988). Although the government would have no trouble effecting amendments in the 'white' chamber of parliament, these would have to be ratified by the 'Indian' and 'coloured' chambers in order to apply to the universities of Durban-Westville and the Western Cape, respectively. Given recent disagreements between President Botha and the leader of the majority party in the 'coloured' chamber, Rev. Hendrickse, de Klerk's amendments may well not have been assured of much support there.

Michael Cherry

Critical report on European fusion 'misses the point'

London

BRITISH fusion researchers have reacted with anger and incredulity to criticisms levelled at the joint European fusion programme, based mainly at Culham in England. A report commissioned by the European Parliament's Scientific and Technological Options Assessment project (STOA) casts doubt on the eventual feasibility of fusion as a viable source of power and implies that the large sums spent on the programme could probably be better deployed elsewhere. "Forecasts about fusion have become couched in an aura of over-optimism. With the resulting risk that serious errors of judgement can be made and, in our view, are being made."

The report, by Colin Sweet, of the London-based Centre for Energy Studies, comes at a particularly sensitive time, with budget wranglings already threatening to delay the fusion programme. The European Commission has yet to approve the proposed fusion budget of 911 million ECU (European currency units; 1 ECU = £0.69) to 1991. Sweet's principal conclusion is that "a full and proper economic

evaluation is overdue" and that such an evaluation be made a condition of the allocation of any resources beyond 1991. The report expresses doubt that the present system for review, by the commission's own fusion directorate, would provide the necessary degree of objectivity. This appears to be about the only point where the programme's supporters are prepared to concur with Sweet.

The politest reaction to the report is that where it gets its facts right it states the obvious but that overall it is essentially misconceived. "We do not consider this a particularly high-quality review," said an official. "There are some points that are valid but I do not think the man really understood the problems." Fusion's proponents say that the programme is being undertaken in the full expectation that it will take decades to overcome the enormous technical hurdles, and that comments in the report belie a fundamental misunderstanding of the programme's purpose.

Simon Hadlington

*Criteria for the Assessment of European Fusion research. STOA project, European Parliament Luxembourg. 10 ECU.

Planet Earth becomes target for International Space Year

Boston

SPACE agency officials from 16 nations and the European Space Agency (ESA) gathered in Durham, New Hampshire, last week to lay the foundation for a mission to Earth as the chief component of International Space Year (ISY) in 1992. (The International Biosphere-Geosphere Program of the International Council of Scientific Unions is separate, but collaboration is planned.)

1992 will be both the thirty-fifth anniversary of International Geophysical Year and the 500th anniversary of Columbus's voyage to America. Given concern over threats to the Earth's environment, the campaign will pay less attention to neighbouring planets and distant galaxies, but will instead concentrate on a better understanding of the Earth, and how the talents of the world's space agencies can be pooled in a coordinated effort to map global environmental change.

Throughout the three-day conference, representatives stressed links between environmental problems: depletion of the

ozone layer, deforestation, desertification, pollution, acid rain and the 'greenhouse effect'. New projects will be in operation by the early 1990s, based on what the conference's working group on 'space-based ISY mission components', a "fantastic array" of orbiting instruments.

The plan is that national space agencies should supply full inventories of scheduled projects for entry into a global compendium. One of the projects likely to figure prominently is NASA's Upper Atmosphere Research Satellite (UARS), to be launched in the early 1990s, whose chief purpose is to determine the extent and durability of the ozone layer. Instrumentation will be provided by the United States, Britain, Canada and France, with international data analysis.

Another project is the Japanese Advanced Earth Observing Satellite which may, among other instruments, carry a 'scatterometer' to measure sea-surface wind velocity. The ESA's remote-sensing satellite mission, ERS-1, and a joint French-US TOPEX/Poseidon pro-

gramme to be launched aboard an Ariane rocket, will both provide a wide array of new data on ocean circulation patterns and variability. Other projects are planned by the Soviet Union and India.

Despite the large number of scheduled launches of instruments that could be included in ISY, the consensus at the meeting is that there is very little time to organize a truly integrated international programme by 1992. Most participants agreed that the more important goal is to establish the foundation for a consistent long-term strategy.

One obstacle to an integrated programme, that of organizing the data, was addressed by several working groups. The full extent of the database was likened by some to "literally hundreds of times that of the US Library of Congress", which is why one working group proposed dividing reference databases into units concerned with different topics such as vegetation patterns, soil moisture, permafrost, sea level, ocean colour, sea surface temperature and stratospheric chemistry.

One clear theme is that, thanks to vast new stores of data from space, the Earth environment must increasingly be seen as a system in which no single variable can be understood in isolation. **Seth Shulman**

Australian in space?

Sydney

At talks in Canberra last week, Vyacheslav Dukov, the vice-chairman of the Soviet space agency, offered a flight into space for an Australian cosmonaut. The offer comes amid greatly increased discussion of joint projects in space stimulated by the signing of the US-Soviet Space Research Cooperation Agreement last December.

Bruce Middleton, executive director of the Australian Space Office, believes a manned flight will not happen before the next century. "If at some future date an Australian experiment needed technical expertise, we would consider sending a scientist into space", he said.

The Soviets are particularly interested in using Australia's proposed commercial launch facility in Cape York, North Queensland, for launching Soviet rockets. According to Ken McCracken, director of the Office of Space Science and Applications of the Commonwealth Scientific and Industrial Organization, Western countries have been reluctant to use Soviet vehicles because of concern over transfer of sensitive technological information. "The Russians are pragmatic", he said.

"They know that, regarding latitude and climate, our facility will be superior to the one at Cape Canaveral. They also know that they will not break through the US embargo on information arising from American technology entering the USSR. We are offering them a commercial alternative."

Tania Ewing

AIDS brochure launched despite doubts

Washington

THE US Public Health Service last week unveiled its educational brochure on AIDS, to be mailed to every household in the United States by the end of June. Over 107 million copies of the eight-page leaflet entitled *Understanding AIDS* will be mailed, at a cost of \$17 million.

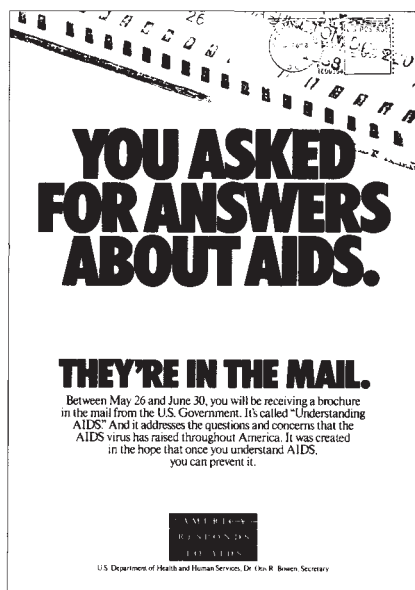
The brochure contains a frank discussion of the modes of AIDS transmission and identifies who should be tested for anti-

bodies to human immunodeficiency virus (HIV), the virus causing AIDS. It gives equal space to discussions of condom use, abstinence from drug use and care in selection of sexual partners to prevent contact with the virus.

The long-awaited brochure should have been mailed last autumn, but was withheld until a compromise could be reached between President Ronald Reagan's approach of instructing people to "just say no" and Surgeon General C. Everett Koop's advice that they should use condoms. The \$20 million set aside for the production and mailing of the brochure in last year's budget was absorbed into an expansion of the national AIDS hotline and the distribution of a less specific brochure to non-profit organizations and state and local health departments (see *Nature* 331, 470; 1988).

The effect of the brochure on the public's knowledge of AIDS will be gauged through responses to the current National Health Interview Survey on a multitude of health topics administered by the National Center for Health Statistics. Questions about AIDS were added to the randomly administered survey starting in August last year. The survey will also attempt to assess the effect of increased knowledge about AIDS on behaviour: questions about sexual practices and drug use are expected to be added this summer.

Carol Ezzell



The poster announcing Koop's message.

Mental illness moves into the genetics lab

Santa Barbara, California

THE National Institute of Mental Health (NIMH) has travelled further down the road from the therapist's couch to the sequencing gel, with a call for grant applications in molecular neurobiology. Underscoring its commitment to molecular approaches to mental illness, it sponsored a three-day conference on molecular neurobiology here last week, designed as a 'mixer' to encourage the flow of ideas between NIMH researchers and molecular biologists.

The NIMH's commitment to mental health has taken many forms over the four decades since its inception, with support for basic research at times overshadowed by service functions. NIMH played a role in the management of community outpatient programmes during the heyday of belief in the 'social cure' in the 1960s.

The current trend towards increased support for basic research has its roots in the formation, in 1981, of the institute's Neurosciences Research Branch. The subsequent recommendation, by panels of leading neuroscientists, of vigorous support for molecular biology, arose out of the recognition that mental illness is likely to have genetic components.

"We're optimistic about the payoff from molecular biology", said Stephen Koslow, head of the Neurosciences Research Branch. Koslow hopes to persuade researchers working on genetic and molecular mechanisms of nerve cell function and development to consider how their systems may illuminate underlying causes and potential treatments of mental illness. There has been much progress in recent years, for example in the cloning of the genes for neurotransmitter and hormone receptor systems, as well as ion channels, and enzymes involved in neurotransmitter synthesis. Within the next decade, NIMH hopes research will reveal connections between genetic errors and various forms of mental illness.

Eric Kandel, of Columbia University, keynote speaker at last week's conference and long-time recipient of NIMH funding, does the sort of work the institute wants to support, said Steven Zalcman, head of the neurobiology programme. Kandel has taken his studies of learning and memory in the sea slug, *Aplysia*, from the level of neural circuitry to analysis of the molecular mechanism of memory storage. In addition to increasing its portfolio of grants to individual molecular biologists, NIMH hopes that its January 1988 *Decade of the Brain* report will "light a fire under Congress", said Zalcman. **Marcia Barinaga**

Shake-up proposed for all four of the UK research councils

London

THE long-running debate about the division of responsibilities between Britain's four scientific research councils has resurfaced, with three of the councils insisting that the time is ripe for change. Only the largest of the four, the Science and Engineering Research Council (SERC), wants to retain the status quo — under the proposals being put forward, it would stand to lose its role in the biological sciences.

The various schemes have been suggested during evidence to a House of Lords select committee that is examining Britain's food and agricultural interests. The most radical proposal has been put forward by the Natural Environment Research Council (NERC), which says that the "previously fairly clear" divisions of responsibility between the different councils has become blurred and confused, "with the likelihood of increasing overlap and duplication". NERC proposes to amalgamate with the Agricultural and Food Research Council (AFRC) to create a Natural Resources Research Council. NERC cites several areas of common interest, such as forestry and land use, to support its case. The new council would retain the councils' biological interests in order to keep "the vital links between the biological and physical sciences". AFRC's solution is to establish a single non-medical biological research council.

The Medical Research Council (MRC) firmly holds that the biological science

research administered by the SERC would be better farmed out to the three 'mission-led' councils. It cites SERC's programmes in plant and animal biochemistry and physiology as well as "many components of its biotechnology programme" as areas which make it "inherently difficult to coordinate initiatives effectively". MRC says that while it successfully developed mechanisms with the AFRC to coordinate activities, attempts to do the same with SERC had failed. MRC is opposed to the establishment of a single biological research council.

SERC makes a spirited defence of the present system. It points out that its annual expenditure on non-clinical biology is £23 million. It supports 1,700 research students in biology, of whom 600 are on industrially sponsored schemes.

Because of its other interests it is best placed to encourage close collaboration in fundamental and strategic areas between the life scientists and the physical and chemical scientists and mathematicians supported through its science board. Its interests in engineering research provide a valuable means of exploiting biology. Current systems for cooperation and coordination between the research councils are sufficient to avoid substantial overlap and duplication, SERC maintains. It will be for the Advisory Board for the Research Councils to try to make some sort of sense of the conflicting schemes.

Simon Hadlington

Switch to PWR could produce skills shortage

London

BRITISH industry and government were warned last week that the renaissance of the nuclear power programme, which until last year had lain dormant for almost a decade, could expose a vacuum in skilled manpower, resulting in increased reliance on imports of power station components and fuel. In an assessment of the industrial impact of the decision to construct a pressurized water reactor (PWR) at Sizewell, in Suffolk, the government's Advisory Council on Science and Technology (ACOST) says that although the decision to embark on a PWR-based programme was not likely to result in a severe immediate shortage of available manpower in research, development or design, if all the power stations were built too close together in time, or if there was a significant, simultaneous coal-fired power station building programme, "there could be difficulties due to an inadequate supply of sufficiently well trained and experienced project

managers and site personnel".

Since permission was granted for the construction of a PWR at Sizewell early last year, the Central Electricity Generating Board (CEGB) has made clear its intention to build a family of identical reactors. ACOST supports CEGB's position that all efforts should now be directed towards the construction of the PWR and that no new advanced gas-cooled reactors should be built.

On the question of Britain's fast reactor research programme, currently the subject of intense debate and speculation, given that the electricity supply industry is soon to be privatized, ACOST notes "with approval" that funding for the project (which totalled £128 million last year) is moving increasingly away from the government to the CEGB and recommends "that this process should be encouraged".

Simon Hadlington

* The Industrial Impact of Sizewell 'B'. ACOST (HMSO, London, £4.90).

Mono Lake radiocarbon puzzle blamed on nuclear waste disposal

Berkeley & Washington

A GROUP of geochemists studying gas exchange in Mono Lake, California, has accidentally uncovered changes in radioactivity levels that suggest the lake may have been the site of clandestine nuclear waste disposal in the 1950s and 1970s.

The research team, headed by Wallace Broecker of Columbia University's Lamont-Doherty Geological Observatory, found that radioactive carbon-14 levels in water samples from the lake jumped 15 per cent in the 1950s and 5 per cent in the 1970s, a total of about 20 curies of radiocarbon. The increase is 15 times too great to be accounted for by fallout from atmospheric nuclear tests in nearby Nevada, says Broecker, and studies in two other Nevada lakes show no such discrepancies. After eliminating every known natural explanation for the jump in

remains that other radioactive elements entered the lake with the radiocarbon, especially if the rise was due to nuclear reactor or weapons waste. Broecker says his team plans to test core samples from the sediments at the bottom of the lake for radioactive caesium, strontium and other elements. Other Lamont-Doherty researchers found slightly elevated levels of plutonium in the lake, which they attributed to a high solubility in the lake's alkaline water. Broecker suggests that their conclusion may need to be reevaluated in light of the carbon-14 discovery.

The percentage of carbon-14 in Mono Lake carbon is naturally low, because hydrothermal vents in the lake are continually adding carbon dioxide that is completely free of carbon-14. Even with the mysteriously added radiocarbon, the carbon-14 to carbon-12 ratio in the lake is less than that in ocean water.

Broecker has been monitoring radiocarbon in the lake for more than 30 years, as part of a study of gas exchange. By 1972 he had noticed that radiocarbon levels in the lake had increased relative to his 1957 measurements very much faster than he expected. A possible explanation was that the lake's unusual characteristics somehow increased the rate at which radiocarbon in the atmosphere, produced by nuclear bomb tests and reactors,

entered the lake waters. But experiments with an inert chemical tracer, SF₆, showed no abnormalities in the rate of gas exchange between lake and atmosphere. And although the high alkalinity of the lake water did enhance chemical uptake of carbon dioxide, it did not do so by nearly enough to explain the increased radiocarbon.

Broecker was able to chart the radiation discrepancy more accurately when he recently compared notes with Scott Stine, a graduate student at the University of California, Berkeley, who had been estimating carbon-14 levels in the lake by sampling the calcium carbonate buildup on pebbles and wood along the steadily receding lake shore. Pebbles collected from the 1919 and 1938–47 shorelines, confirmed that there had been a massive increase in radiocarbon sometime after the Second World War, strongly suggestive of clandestine disposal of radioactive waste.

Broecker's first guess was that "the graphite from the World War II Univer-

sity of Chicago graphite reactor" was a candidate. Some uncertainty does surround the fate of Chicago Pile Number One (CP1), with which Enrico Fermi first demonstrated that a nuclear chain reaction could take place. Its 750,000 pounds of pure graphite was moved out to the site of what is now Argonne National Laboratory in 1943. Some small (and non-radioactive) pieces of the graphite ended up, encased in lucite, as souvenirs for those who worked on the project. Old-timers at Argonne are not certain what happened to the rest, but believe that it went for reprocessing to Oak Ridge National Laboratory sometime in the 1950s. A Department of Energy spokes-

Danger from Mono Lake water level

Washington

A separate report, commissioned by the California Department of Fishing and Game from the University of California, Santa Barbara, charts the future of Mono Lake.

Since 1940, Los Angeles has been taking some 17 per cent of its water needs from the streams that flow into Mono Lake. As a result, the level of the lake has fallen 45 feet, its volume has halved and its salinity has doubled.

If water continues to be taken at the present rate, many of its distinctive tufa towers and substantial parts of the gull habitat will be endangered by 1991. By 2000, bird life will disappear and later in the next century Mono Lake will become another Dead Sea. Los Angeles will need to cut its water imports from the area by almost half to keep the lake as it is.

Alun Anderson

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The distinctive tufa towers in Mono Lake.

radiocarbon, Broecker finds himself left with the possibility that clandestine dumping of nuclear waste from government weapons tests is to blame.

Mono Lake, a highly saline lake with no outlet, has been the subject of controversy over its drop in water level, due to diversion of its feeder streams to provide drinking water for Los Angeles. The lake, which has a salinity more than twice that of sea water, and a pH of 10, has no fish, but its algae, brine flies and brine shrimp provide food for large populations of migrating birds. The lake is known for its calcium carbonate formations, called tufa towers, formed where calcium-rich spring water enters the carbonate-laden lake.

The 20 curies of radiocarbon estimated to have been added to the lake increased total radiocarbon by only 20 per cent, because the lake water is so carbon-rich, says Broecker. He stresses that the additional radiocarbon poses no threat to the lake ecology, and certainly not to humans, as the lake is used neither for drinking water nor recreation. But the possibility

man dismissed the possibility that the Chicago graphite could have been transported across to Mono Lake as "ludicrous".

Additional data analysis by Broecker and his team has in any case shifted the first influx to a later date than first thought, between 1952 and 1957. By that time numerous bomb and reactor programmes were under way.

Military tests by the US Navy were even carried out in Mono Lake itself during the 1950s. But residents around the lake at the time were told the explosive tests were non-nuclear in nature, according to Emily Strauss, of the Mono Lake Committee. Navy and weapons laboratory officials have denied any responsibility for the elevated radiocarbon levels. Broecker hopes that publishing his results may encourage someone to come forward with an explanation — natural or unnatural — for the mystery.

Marcia Barinaga & Alun Anderson

The thesis that won't go away

SIR—I sympathize greatly with Beverly Halstead's position on PhD theses (*Nature* 331, 497; 1988) and almost entirely disagree with the conclusions.

For one thing, experience on a higher degrees committee has taught me that requirements of different subjects are very different, and it is possible that detailed requirements for one subject will not suit another, even when these subjects are closely related.

The general principles, of course, remain the same. These are to demonstrate that the student knows how to initiate, undertake and complete a piece of research and how to present it, evaluate its significance, propose further development and continuation and how to profit from negative results obtained in the course of the work.

To finish a thesis in a reasonable length of time — and in chemistry three years from the first degree is quite enough — it may be necessary to relax the requirement for completion of a piece of work; and in this case adequate discussion of negative results is more than necessary to show the author has profited from them.

I do not agree that theses are never referred to again after examination. In my experience it is not usually difficult to obtain photocopies, and again in chemistry, these often contain experimental

details missing in publication which may be very important in repeating or extending the work.

D. A. H. TAYLOR

*Department of Chemistry &
Applied Chemistry,
University of Natal,
King George V Avenue,
Durban 4001, South Africa*

SIR—Although I am not familiar with the details of the individual case reported and personally experienced by Beverley Halstead in which an established scientist with a distinguished research record was denied a PhD degree, in the extract from the external examiner's report there is clear evidence that the work as submitted by the candidate was taken seriously; and I understand that he was also given the opportunity to decline the MPhil. degree and try again with an enlarged standard-type version of the original thesis. In other words, according to the examiner's report, it would have been perfectly acceptable if the published text were directly incorporated into the thesis, which is something that not all universities allow. To me the external examiner's point of view does not seem unreasonable. In fact, if I were in his position and if my role in the process was reduced to listening to some verbal comments on work which

at that point had been reviewed by scores of referees, then I sense that I would probably have considered this whole exercise as degrading or, at best, as a waste of time. In connection with this particular case, Halstead also pointed out that PhD candidates are frequently judged on the basis of their output as measured in terms of the number of pages in their theses. I am more than willing to believe him in this respect and I find that this particular attitude would cause outgrowths of nightmarish proportions if in the future the number of published pages were to play an analogous role.

PETER SENN

*Kistlerstrasse 32,
8364 Reichenburg, Switzerland*

India and China

SIR—The letter by Upinder Fotadar on "India and China" (*Nature* 332, 390; 1988) missed almost all points.

First, Fotadar believes that available data in the West about the development of the communist countries are usually inaccurate and insufficient because of the closed-door policy of those countries. His conclusion, therefore, based on the available data, that India leads China in most fields, is incorrect. India leads China in some fields.

Second, it is true that India has the third largest technical manpower in the world. But the advancement of a country depends on the quality of its manpower rather than the quantity, as can be seen in Great Britain, Japan and West Germany.

Third, the scientific achievement of the Eastern-bloc countries cannot be judged solely on the number of research papers published in the Western journals. Statistically speaking, there are more articles in *Nature* from India than from the Soviet Union. Moreover, Japan, a Western country, produces an enormous number of research papers. But Japanese papers published in journals such as *Nature* and *Science* are relatively few because of the lack of manpower to translate them. Is Japan behind India?

Fourth, Fotadar falsely says that China was defeated in the Sino-Vietnamese conflict in 1979. The truth is that China won the war, but it paid the price. That is the nature of war. Moreover, if China had been defeated, how about the wars fought between Russian and Finnish, Chinese and American in Korea, Chinese and Indian in Tibet, Vietnamese and American? Did these wars reflect the backwardness of Russian, American and Indian?

JAY J. TU

*Program in Advanced
Ceramic Heat Engine,
Department of Materials Engineering,
Virginia Polytechnic Institute
and State University,
Blacksburg, Virginia 24061, USA*

Radiocarbon-dating the shroud

SIR—Dr M.S. Tite (*Nature* 332, 482; 1988) has revealed the new procedures decreed by Cardinal Ballestrero, Archbishop of Turin, for radiocarbon-dating the Shroud of Turin. These differ so remarkably from those of the original protocol agreed by all parties at the Turin Workshop held in the fall of 1986 and chaired by Professor Carlos Chagas, president of the Pontifical Academy of Sciences [H.E. Gove, *Nucl. Instr. Meth. Phys. Res. B* 29, 193 (1987)], that a brief comparison of the two seems in order.

1. The involvement of seven laboratories has been reduced to three. This eliminates the possibility of detecting a mistake made in the measurement by one or more of the three laboratories. As Tite knows, such mistakes are not unusual.

2. The use of both decay counting and accelerator mass spectrometry (AMS) has been changed to AMS only. The two methods are distinct and independent.

3. The amount of cloth each AMS laboratory receives has been increased by almost a factor of two. With this much material, several more laboratories could have been included.

4. Representatives of the three laboratories will not be permitted to observe the sample removal from the shroud. Tite will

be the only independent scientist present at this operation.

5. The shroud and control samples will not be unravelled and thus, despite Tite's comments to the contrary, the shroud sample will be much more easily identifiable.

6. The scientific body connected with the Roman Catholic church which has a high reputation in the world of science, the Pontifical Academy of Sciences, has unaccountably been excluded from official participation in any aspect of this important and controversial radiocarbon measurement.

7. The acknowledged textile expert selected at the Turin Workshop to remove the shroud sample has been replaced by some unnamed person.

All these unnecessary and unexplained changes unilaterally dictated by the Archbishop of Turin will produce an age for the Turin Shroud which will be vastly less credible than that which could have been obtained if the original Turin Workshop protocol had been followed. Perhaps that is just what the Turin authorities intend.

H.E. GOVE

*Nuclear Structure Research Laboratory,
University of Rochester,
Rochester, New York 14627, USA*

The big bad book of the 1930s

People should re-read long-forgotten books from time to time, partly to understand how they may have been led astray but also for mere old time's sake.

THE 1930s were the decade of the big book. Everybody has a handful of titles tucked into the back of his mind — Born's *Atomic Physics*, Dirac's *Quantum Theory*, Sidgwick's *Organic chemistry of nitrogen* and books like that. One does not have to have been alive at the time to have known of them. But the big book of the 1930s that hardly anybody remembers, and which fewer have read, is in many ways the most interesting of them all, if only because it is a touching reminder of how it is possible to be at once dead right and dead wrong, but entirely oblivious of the second possibility. The book is *Relativity theory of protons and electrons* by A.S. ("Sir Arthur") Eddington, published by Cambridge University Press in 1936. Technically a summary and synthesis of Eddington's research on the connection between quantum theory and relativity in the preceding eight years, the book comes out as a kind of declaration of faith that the problem of the Universe is now solved, but that there may be a little tedious computation left to be carried out.

Much of that impression has to do with the style of Eddington's prose, so pure in its self-confidence that even admissions of previous error read assertively, as when Eddington lists ten previously published papers in the *Proc. R. Soc.* and continues "...it was suggested that ... correspond to electrons of opposite spin ... this misjudgement persisted ... I think there is no other point on which I went so completely astray ... the other lines of development..., though sometimes requiring substantial amendment,... contain advances which in principle have been retained."

But the essence of the appeal of Eddington's book is the sheer brazenness of its conviction that the problem (of the Universe, of particles and all that) had been solved. "I think it will be found that the theory is purely deductive, being based on epistemological principles, and not on physical hypotheses", says the *Introduction*. The last sentence of this moving text, 329 pages later, is "and so we may look forward with undiminished enthusiasm to learning what lies hidden in the atomic nucleus — even though we suspect it is hidden there by ourselves." If Jean-Paul Satre had been known in 1936, Eddington would have been called an existentialist.

In the event, he was called an aristotelian by the late Sir Herbert Dingle (who

separately, but successfully, sued *Nature* for libel in 1973) and found the label insupportable. Dingle, a polemicist among academics, who gave *Nature's* readers the first news of Eddington's big book (139, 784; 1973), seems to have been the one to coin the word "cosmolatry" (between cosmology and idolatry) to describe the faith of those who take the Universe to be their god. The know-all philosophy, said Dingle, has "grown so confident that we have no need to unmask it; it has itself discarded masks..." And (of Eddington), "if he had a philosophy big enough for his intuitions, we should know something".

Eddington naturally replied (*Nature* 139, 1000; 1937) in the manner of a British civil servant by describing Dingle's attack as "entertaining", but surprisingly continued by reaffirming his belief in the *a priori* Universe. "One consequence of the new outlook which, I believe, rouses the passions of the Galileans, is that a law of nature has a compulsory character, and is not merely an empirical regularity which we have hitherto found to be fulfilled in our limited experience."

Schrodinger, when he came to review the book almost a year later (*Nature* 140, 742; 1937), did not give his hand away on these grand philosophical issues, but stayed with the text. Thus "...this theory has succeeded in accounting for bewilderingly numerous and complicated facts from comparatively simple (though strange) first principles... we have before us a sketch of unusual grandeur, of which not the details alone need further development and, maybe, further modification".

A further byplay (also recorded in *Nature*) is relevant in the sequel. One of the other big books of the 1930s was E.A. Milne's *Relativity, gravitation and world-structure*. Milne's book was reviewed in *Nature* anonymously by Eddington, who savaged it in the politest possible way: "when he finds... particles the properties of which are like those of cosmic rays, we congratulate him on the happy resemblance of his selected model to the actual universe. But the particles... occur in his model because he put them there. That he did so undesignedly does not affect the question." Ironically, Milne (who was the nicest of men) was later that same year proposing the vote of thanks at a lecture by Eddington on the structure of the Sun.

So what did Eddington's own book accomplish? In his grand design, the scheme was to calculate everything from nothing. Much in sneaking awe of Dirac, Eddington was similarly preoccupied by the meaning of the numbers that seemed to dominate the structure of the Universe, not merely the fine structure constant ($1/137$) but the large numbers which are the ratios of the coupling constants for the forces between particles. His book (well worth reading even now) contains a calculation of the ratio of the mass of the proton (calculated numerically as 1847.6) and of the number of particles (protons and electrons) in the Universe, roughly 10^{79} .

Milne, who returned to Oxford as soon as the Second World War was over, used to hold a seminar twice a week on topics of his audience's choice. Because there were only a handful of people in the audience, most people were required to prepare something to talk about quite often. Foolishly knowing nothing of the affair of Eddington's review, I volunteered on one occasion to talk about Eddington's book. Milne was exceedingly kind, allowing that it was a difficult conception the grand old man had set himself, and that it was natural that he should have converted young people so easily to his way of thinking. He said that Eddington was devoid of physics.

That must be so. Certainly he had the bad luck to be publishing his big book when the whole of physics was changing in ways that he could not have guessed (but there are signs, in the later chapters, that he was trying to adapt to the discovery of the neutron four years earlier). But Eddington's book also has some entertaining ideas that are worth knowing about even at this late stage.

For example, it is a good way of learning about Cayley algebra; there is a discussion of the reasons why the simplest approximations to the real Universe must have five dimensions (one for time and one for curvature) and why, in general, there must be ten (shades of superstrings, but also a simple count of the number of independent variables in a symmetrical metric). Getting from 10 (or 16) to 136, to approximate to the fine-structure constant, requires a little more incredulity — and then adding one to make 137 is an assault on common sense. But who, among the aristotellians, will think that bad form?

John Maddox

Archaeology

Origins of the English castle

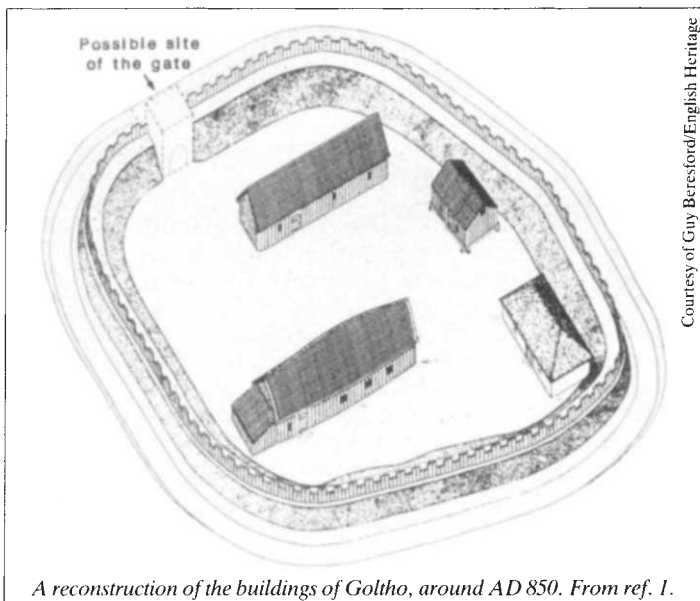
Richard Hodges

THE recent publication¹ of Guy Beresford's excavations of the small castle at Goltho in the east of England provides an opportunity to assess the warning² by the historian R.A. Brown about the problems inherent in elucidating the origins of castles. Brown believed that archaeologists' ignorance of history (and vice versa), together with the lack of any particular principle of private fortification in England before the Norman conquest in 1066, could call into question any conclusion reached by archaeologists about how castles originated.

Goltho is 14 km east of the town of Lincoln in low-lying claylands. Beresford excavated not only the small 'motte and bailey' castle, but also the deserted mediaeval village (which he has described elsewhere³). In his new report¹, he describes a remarkable sequence of earlier occupation beneath the Norman castle levels, which he defines in periods. Period 1 consisted of remains of a Romano-British village belonging to about AD 50–200. Period 2 is a rare example of a Middle Saxon village, from about 800–850, of which parts belonging to at least three farms were exposed. Period 3 is a crucial stage in the history of the site. It comprises traces of a bow-shaped hall, weaving sheds and other buildings, all set within low fortifications. Beresford believes that this period occurred in about 850–950 and describes it as a fortified manor — in a sense, the earliest known English castle (see figure).

Period 4, about 950–1000, is a grander version of the preceding complex, comprising traces of a large hall, weaving sheds, a bower and a kitchen. Period 5, 1000–1080, consists of remains of a substantial manor built around three sides of a courtyard. In period 6 (1080–1150) a small mound (motte) was erected over the earlier buildings, on top of which was constructed a tower, with a 'cramped' building alongside in the fortified enclosure (bailey). In about 1150, Beresford's period 7, a substantial aisled hall was built on the mound, replacing the tower, which is thought to illustrate the "great prosperity of its builder". Shortly afterwards, however, with the reassertion of royal power by the Angevin kings, castles like Goltho were not granted licences, and were deserted in favour of more spacious unfortified manors.

Goltho is not the first private fortified site in England to be attributed to the Anglo-Saxon period. Excavations of the earthen ringwork at Sulgrave, Northamptonshire, for instance, show⁴ that this Norman site had origins in the pre-conquest period. Brown², as a result, has had to come to terms with some irrefutable archaeological facts. But Goltho is the first of these sites to be published in a definitive fashion, permitting the excavated information to be evaluated. Further-



A reconstruction of the buildings of Goltho, around AD 850. From ref. 1.

more, until now no-one imagined that there were Anglo-Saxon castles in pre-Viking, ninth century England. There is therefore some merit in scrutinizing more closely the evidence for Beresford's claims for a fortified manor at Goltho in 850–900.

The remains of the Middle Saxon village attributed to period 2 are remarkably consistent with other sites excavated in the Midlands of England, of which the pre-eminent is at Raunds, Northamptonshire⁵. But the transformation of such a site as early as 850 into a fortified manor belonging to "a man of considerable regional importance" is without any precedent.

Has Beresford dated period 3 correctly? His dating depends mainly on three criteria and secondarily on the finds associated with the buildings. The three primary features are: (1) no evidence that Goltho was "ravaged by Viking raiders" when they overran the area in the late 860s and early 870s; (2) the construction of the fine bow-shaped hall is best paralleled by the West Saxon royal hall found at Cheddar and belonging to the late ninth

century — the carpentry and design, Beresford contends, is very unlike that known from Denmark, thus suggesting this to be pre-Viking workmanship; and (3) he found no artefacts of Scandinavian origin — indeed, the name of Goltho, first recorded as *Golthawe* in 1220–35, is of Old English rather than Scandinavian derivation.

The secondary criteria for dating period 3 are the artefacts that were associated with the buildings. Beresford found a small range of potsherds (broken pottery) but no coins. These pots can be compared with those found in recent excavations of a workshop area and in a pottery kiln in Lincoln. These excavations have the benefit of many coins, so the archaeological features and associated pottery can be dated with comparative confidence to the early to mid tenth century^{6,7}, in fairly good agreement with Beresford's chronology.

Beresford's work fails to take account of Brown's warning, and in particular he follows the traditional assumption that Scandinavian place-names in England mean that there were many Danish settlers. He assumes that the Danish invasion of the 860s left its mark on small farming communities, and that the settlers brought artefacts that distinguish them ethnically from the Anglo-Saxons. Such historical criteria may not be correct. It is now being realized that the Danes concentrated on attacking elite settlements, migrated in small numbers, and contributed significantly to the development of Anglian culture after conquering these territories as well as introducing Frankish crafts⁸. In such circumstances a more accurate date for the fortified manor would be obtained using pottery rather than relying on historical reasoning.

The pottery strongly suggests that the Anglo-Saxon manor at Goltho dates to the tenth century and may be one example of the many changes in English society as a state replaced earlier tribal political systems⁹. The English fortified manor may

1. Beresford, G. *Goltho. The development of an early mediæval manor, c.850–1150* (HMSO, London, 1987).
2. Brown, R.A. *Archaeol. J.* **126**, 131–148 (1969).
3. Beresford, G. *The mediæval clay land village: excavations at Goltho and Barton Blount* (Soc. Mediæval Archaeol., London, 1975).
4. Davison, B.K. *Archaeol. J.* **134**, 105–114 (1977).
5. Cadman, G.E. *Mediev. Archaeol.* **27**, 107–122 (1983).
6. Adams, L. *Mediæval pottery from Broadgate East, Lincoln, 1973* (Council for British Archaeology, London, 1977).
7. Blackburn, M. et al. *Early Mediæval Coins from Lincoln and its Shire, c.770–1100* (Council for British Archaeology, London 1983).
8. Smyth, A.P. *Scandinavian Kings in the British Isles, 850–880* (Clarendon, Oxford, 1977).
9. Hodges, R. *Dark Age Economics* (Duckworth, London, 1982).
10. Hodges, R. *Nature* **321**, 109 (1986).

owe its origins ultimately to the small private fortifications that first occurred in the Frankish kingdoms in Europe in the later ninth century. But it would be wrong to assume that private fortifications like Goltho are an index of Anglo-Saxon feudalism, as Brown implies. The archaeology of English towns like Lincoln and villages like Raunds reveals another social

pattern, and shows that England was a wealthy country on the eve of the Norman Conquest¹⁰. □

Richard Hodges, Department of Archaeology and Prehistory, Sheffield University, Sheffield S10 2TN, is visiting professor at the Forhistorisk-Arkaologisk Institut, Københavns Universitet, 1467 København-K, Denmark.

Palaeontology

Faunas of a southern world

Leigh M. Van Valen

THE mammals of South America look exotic from a northern perspective — ant-eaters and other edentates, camels called llamas, diverse endemic groups of monkeys, rodents and bats, and so on. It has become reasonably common knowledge that the faunas were even more different before the Pliocene joining of the Americas. But in both areas the mammals were all placentals and marsupials, and comparable similarity has been more or less assumed for earlier times. Two remarkable faunas recently described¹⁻⁶ by José Bonaparte at the Museo Argentino de Ciencias Naturales, Buenos Aires, and his collaborators, show that this was not the case.

In the Cretaceous, the mammals of North America consisted mostly of tribosphenids (marsupials and placentals and their immediate forebears) and multituberculates, with a very few rare stragglers from earlier groups. Asia was similar, with more placentals and no marsupials. But in South America we now see a complete absence of tribosphenids and only a very divergent group of multituberculates (if that is what they really are). Multituberculates were the longest-lived order of mammals, vaguely rodent-like in adaptation but part of the pre-tribosphenid basal radiation of mammals.

The bulk of the Cretaceous mammals in South America were of groups which had flourished to the north in the Jurassic: triconodonts, symmetrodonts, dryolestans and amphitherians. These were also part of the basal radiation of mammals and are known mostly from jaws and teeth. The last three groups belong to the pantotheres, the group ancestral to tribosphenids. Most of the families in the South American Cretaceous are unknown elsewhere. The dryolestan *Mesungulatum* partly converged on primitive ungulates (hoofed mammals), and it was even thought to be an ungulate until lower teeth were found. A peramurid-grade amphitherian, *Vincelestes*, a divergent member of the derived group immediately ancestral to tribosphenids, is known from six skulls (the only pantothere skulls known anywhere) and postcranial material still

undescribed. I have seen much of the material and agree with Bonaparte's interpretations for the most part.

The most extraordinary form, though, is *Gondwanatherium*, which has quite high-crowned (hypsodont) molars. No other Mesozoic mammal was hypsodont, even those in Mongolia, which was arid then as now and where the animals undoubtedly had grit blowing on their food. There were no pampas of grasses with opal phytoliths to wear down teeth; the nature of the ground cover is unknown.

Bonaparte thinks that *Gondwanatherium* was an ancestral edentate because of its hypsodont teeth and geographical location. Its hypsodonty, though, is clearly formed entirely by an upgrowth of cusps and of the lower part of the crown, whereas that of edentates seems to be formed by an elongation of the root. Thus, the similarity is probably non-homologous, not a result of continuity of information. Also, loss of enamel would be dysfunctional for an animal with hypsodont teeth. Indeed, much of the evolution of the edentates is an attempt to compensate for the early and apparently triphyletic⁷ loss of enamel in that group. I do not know what *Gondwanatherium* is. Its own order, *Gondwanatheria*⁸, does seem appropriate, though, even if proposed in a fit of oversplitting. Possibly it is even a hypsodont monotreme; its crown pattern has vague resemblances to the early Cretaceous monotreme *Steropodon* of Australia⁹.

The better-known fauna, of the Los Alamitos Formation, is late Cretaceous, near the Campanian-Maestrichtian boundary. The other, from the La Amarga Formation, is earlier, about Hauterivian. Their overall aspect is similar and suggests some generality for the southern fauna. Many reptiles and birds are also of endemic groups^{3,10}, from these formations and elsewhere.

We know nothing of the Cretaceous mammals of the other parts of the disintegrating Gondwanaland except for *Steropodon*. As Bonaparte notes, the southern fauna of the Cretaceous may

have been more widespread. The reptiles do indicate this; perhaps Africa and India will eventually disclose their own Gondwanatherian mammals. And did primitive therians such as pantotheres reach Australia only to be done in by others, even monotremes?

The Gondwanatherian Fauna did not survive the Mesozoic. One genus, *Sudamerica*, is known from the later Palaeocene¹¹. Two earlier Palaeocene¹² faunas (described as Cretaceous¹³) lack any trace, and this is the case after the Palaeocene also. Some of the non-mammalian elements of the fauna did survive, however³. It is even conceivable that the enigmatic Miocene genus *Necrolestes*¹⁴, usually thought to be a marsupial, is a late-surviving Gondwanatherian pantothere. The contemporaneity, precise timing and causes of the extinction remain unknown.

The earlier Palaeocene faunas are interesting in themselves. In addition to the expected precursors of later groups and the possible survival of dinosaurs, there are representatives of two orders of placentals which are otherwise unknown in South America except for a minor Pleistocene incursion of one. The Insectivora and Pantodonta each had appreciable success in Holarctica. That they were rapidly exterminated in the south, as was a smaller group also successful in the north, is another blow to the view that isolated biotas are necessarily inferior adaptively.

The immigration of marsupials and placentals occurred in the early Palaeocene¹², as part of a larger incursion (and excursion). Together with a more prolonged and sporadic exchange in the last 10 million years or so of the Cretaceous^{3,10}, this provides an as yet unappreciated datum for palaeogeographers. It also sets a lower time limit on the autochthonous radiation of marsupials in Australia. That the Cretaceous part of the interchange seems to have involved no mammals is not yet explained. □

1. Bonaparte, J.F. & Pascual, R. *IV Cong. Latinoamericano Paleont.* 1, 361-378 (1987).
2. Bonaparte, J.F. & Rougier, G. *IV Cong. Latinoamericano Paleont.* 1, 343-359 (1987).
3. Bonaparte, J.F. *Acta IV Cong. Argentino Paleont. Biostat.* 2, 63-95 (1986).
4. Bonaparte, J.F. *J. vert. Paleont.* 6, 264-270 (1986).
5. Bonaparte, J.F. *Actas IV Cong. Argentino Paleont. Biostat.* 2, 48-61 (1986).
6. Bonaparte, J.F. & Soria, M. *Ameghiniana* 21, 177 (1985).
7. Simpson, G.G. *Am. Mus. Novitates* 567, 1-4 (1932).
8. Mones, A. *Comunic. Paleont. Mus. Hist. Nat. Montevideo* 1, 237-240 (1987).
9. Archer, M., Flannery, T.F., Ritchie, A. & Molnar, R.E. *Nature* 318, 363-366 (1985).
10. Bonaparte, J.F. & Kielan-Jaworowska, Z. *Fourth Symp. Mesozoic Terrest. Ecosyst., Short Pap.* 24-29 (1987).
11. Scillato-Yané, G.J. & Pascual, R. *Ameghiniana* 21, 173-176 (1985).
12. Van Valen, L.M. *Evol. Monog.* 10, 1-79 (1988).
13. Marshall, L.G. & de Muizon, C. *Natn. geog. Res.* 4, 23-55 (1988).
14. Patterson, B. *Harvard Univ., Mus. Compur. Zool., Brevis* 94, 1-14 (1958).

Leigh M. Van Valen is in the Department of Biology, University of Chicago, Chicago, Illinois 60637, USA.

Photosynthesis

Electron-transfer theory in question

Jim Barber

THE reactions by which light is captured and stored as chemical energy in photosynthesis are distinguished by the extremely high quantum yield of the primary conversion process. On page 190 of this issue¹, Fleming *et al.* present data which provide new insights into this process. Using ultrafast spectroscopy, these authors have been able to detect the rate of formation of the initial charged transfer state in the isolated reaction centres of purple photosynthetic bacteria. This occurs in just a few picoseconds and, surprisingly, increases as the temperature is lowered to that of liquid helium.

Their results emphasize that a complete theoretical description of this remarkable primary reaction will require a better understanding of how the protein environment around the reactants influences the very fast electron-transfer rate. One particular challenge which emerges from their observations is to explain, in terms of quantum mechanics, how an intermediate bacteriochlorophyll molecule aids the primary charge separation.

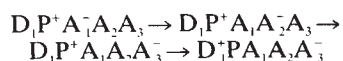
Reaction centre

All photosynthetic organisms have similar molecular machinery for catalysing the primary storage processes of photosynthesis. The machinery is contained within a protein complex embedded in a membrane and called a reaction centre. The reaction centre is excited by the energy of a photon which is initially absorbed by an antenna system of many hundreds of pigment molecules. The energy of the absorbed photon is rapidly transferred by exciton migration to the reaction centre which acts as a trap. All types of reaction centres contain either chlorophyll (oxygenic organisms, such as higher plants and algae) or bacteriochlorophyll (anoxygenic organisms, such as photosynthetic bacteria) which acts as a primary electron donor, P. When excited, P rapidly gives an electron to a primary electron acceptor, A₁, and a charged transfer state or radical pair, P⁺A₁⁻, is created.

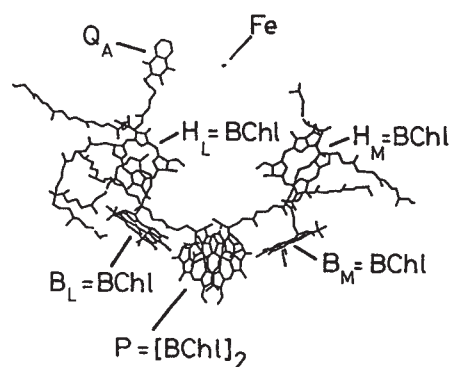


The transfer of an electron from P^{*} to A₁ must be very fast to avoid de-excitation by other routes such as by fluorescence. Also, for the photosynthetic reaction centre to have a very high quantum yield of energy conversion, there must be a negligible back-reaction between P⁺ and A₁⁻. The latter is achieved by rapidly transferring

the electron away from A₁⁻, to secondary electron acceptors A₂, A₃ and so on, and by passing an electron to P⁺ from the secondary donors D₁ and so on.



For some years now, the best experimental system to study these primary electron-transfer processes has been reaction centres isolated from purple photosynthetic bacteria. The importance



Organization of the prosthetic groups within the reaction centre of *Rps. viridis* (from ref. 2). The 2-fold symmetry axis is aligned vertically in the plane of the paper. Note that, unlike the reaction centre structure of *Rb. sphaeroides*³, the secondary quinone Q_B is not present but its position is symmetrically related, around the 2-fold axis, to the Q_A site.

of this experimental system has been particularly elevated by the recent determination of the three-dimensional structure, to a resolution of atomic distance, of the reaction centres of *Rhodospseudomonas viridis*² and *Rhodobacter sphaeroides*³. These outstanding studies indicate that there are two different symmetrically related possible pathways for the primary electron transfer to occur, the L and M branches (see figure).

For reasons which are not fully understood, the pathway of choice is L. The primary donor P is a dimeric form of bacteriochlorophyll (BChl)₂ and is shared by both pathways. In addition to this, there are two monomeric bacteriochlorophylls (B_L and B_M) and two bacteriopheophytins (H_L and H_M). The secondary electron acceptors are bound quinone molecules called Q_A and Q_B. The donor (D₁) to P⁺ is the haem of a cytochrome which may be bound to the reaction centre,

as in the case of *Rps. viridis*, or may be a soluble protein, as in the case of *Rb. sphaeroides*.

According to the crystal structure (see figure), it would be logical if A₁ was B_L, but the data of Fleming *et al.* in this issue¹ do not support this idea. Using stimulated emission in the infrared to monitor the PA₁ to P⁺A₁⁻ transition in both *Rps. viridis* and *Rb. sphaeroides*, they conclude that A₁ is the bacteriopheophytin H_L. According to their data, if there is an involvement of B_L in the electron exchange between P and H_L, then it must include an adiabatic step or be incorporated into a superexchange mechanism with very strong quantum mechanical coupling. Relevant to these possibilities is the finding that rate of primary charge separation increases with decreasing temperatures. At room temperature, the rate for P⁺A₁⁻ formation is identical for both bacterial systems, but at 8 K not only does the rate increase dramatically but it is almost twice as fast in *Rps. viridis* as it is in *Rb. sphaeroides*.

Discrepancy

Although Fleming *et al.*¹ try hard to explain these data in terms of conventional theory for non-adiabatic electron transfer, they are forced to conclude that the validity of the arguments is not convincing, particularly in the case of *Rps. viridis*. The problem is that the components involved in primary charge separation are embedded in a protein environment^{3,4}, and it is this complication which probably underlies the discrepancy between theory and the differences obtained with the two bacterial systems. Indeed, the crystallographic data indicate small differences in organization of the chromophore within the *Rps. viridis* and *Rb. sphaeroides* reaction centres^{3,4}, the implications of which are not fully understood in terms of the kinetics of primary charge separation. Also, the static structure obtained by X-ray diffraction does not reveal the conformational changes that may occur when the reaction centre is activated by light.

Clearly, Fleming *et al.* provide the necessary experimental data which, together with the crystallographic structures, will allow theoreticians to modify existing ideas and to develop new ones about how photosynthetic organisms so efficiently bring about the fastest bimolecular reactions known in photochemistry. □

1. Fleming, G. R., Martin, J. L. & Breton, J. *Nature* **333**, 190–192 (1988).
2. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *J. molec. Biol.* **180**, 385–398 (1984).
3. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H. & Rees, D. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5730–5734 (1987).
4. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *Nature* **318**, 618–624 (1985).

Jim Barber is Professor of Plant Physiology and Head of the AFRC Photosynthesis Research Group, Imperial College of Science and Technology, London SW7 2BB, UK.

Ocean Drilling Program

Plutonic rocks in fracture zones

*Leg 118 shipboard scientific party**

THE main objective of Leg 118 was to sample a section of the lower oceanic crust and upper mantle in the Atlantis II fracture zone on the south-west Indian ridge (Fig. 1). For this, a guidebase has been developed to begin a hole on a hard rock outcrop (see the report of Leg 106 in *Nature* 321, 14–15; 1986). The Atlantis II fracture zone was chosen because the crust there is generally thin and the feature shows great topographical relief, ensuring that lower crustal and upper mantle rocks would be exposed. During Leg 118, we continuously cored a 500.7-m thick section of gabbro at site 735 in the east wall of the fracture zone. The recovered core provides a unique opportunity to study the structure, composition and evolution of lower ocean crust and to follow magma-chamber evolution at a slow-spreading ridge.

During the first half of Leg 118, we tried drilling at several sites (732–734) on a median ridge and on the steep (up to 45°) walls of the fracture zone. Ubiquitous rubble prevented the stable drilling conditions required to drill a deep hole. Site 735 (the principal site) is on a shoal platform in about 700 m of water on the east rim of the transform (Fig. 2). This platform, about 9 km long in a north–south direction and 2 km wide, is one of a series of uplifted blocks which lie on a ridge parallel to the transform. It is remarkably flat, suggesting that it had been eroded to sea level, subsequently subsiding to its present depth. Its position in the magnetic anomaly pattern suggests a crustal age of about 12 million years. A television/sonar survey reveals massive, layered basement outcrops only partially mantled with thin sediment.

A total of 435 m igneous and metamorphic rock was recovered from the 500.7 m cored. The predominantly gabbroic rocks recovered have undergone varying degrees of plastic and brittle deformation, and many have well-developed foliations. Six principal lithostratigraphical units can be recognized by their igneous mineralogy, mineral composi-

tions and metamorphism. The uppermost unit is a 39.5-m-thick sequence of foliated metagabbro with porphyroclastic to mylonitic textures. This sequence is underlain by 140.5 m of olivine and olivine-bearing gabbro (unit 2) followed by 44 m of olivine gabbro containing intervals of iron–titanium oxide gabbro (unit 3). Unit 4 consists of relatively evolved magnetite–ilmenite gabbro containing unusually high amounts of iron–titanium oxides. These rocks contain up to 30 per cent iron oxide and 9 per cent TiO₂ by weight and on land would be suitable for mining. Unit 5 consists of 102.5 m of relatively uniform olivine gabbro and is underlain by a 126-m sequence of olivine-rich gabbros containing layers of troctolite (unit 6). The troctolites are highly magnesian, fine-grained rocks derived from highly mafic magmas.

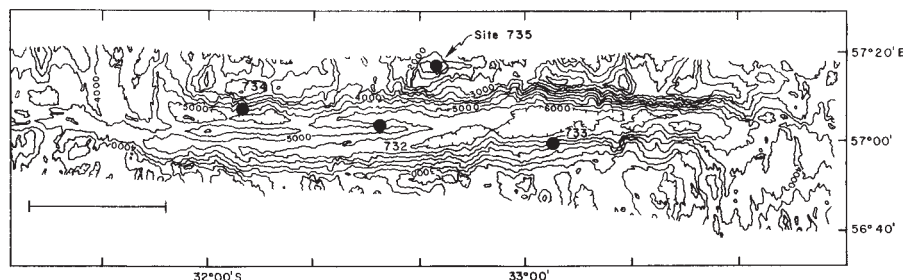


Fig. 2 Bathymetric map (contours of 1,000 m) of the Atlantis II fracture zone on the south-west Indian ridge. Scale bar, 50 km. (Survey by H. Dick, D. Gallo and R. Tyce, 1986.)

None of the contacts between units is clearly magmatic, and two are definitely tectonic. The mean dips of veins and foliation decrease from about 30° at the top to 20° at the base. The lineation defined by mineral elongations is mostly down-dip, indicating that movement during deformation was normal or reverse along the dip direction, rather than strike slip.

An early stage of metamorphism produced highly foliated porphyroclastic, gneissic and mylonitic rocks in some intervals. This high-temperature event was followed by static alteration manifested by numerous veins, 0.5–20 mm wide, filled largely with hornblende and sodic plagioclase and joined by epidote, sphene and chlorite in permeable brecciated zones. In undeformed gabbros, static alteration resulted in development of coronas of secondary minerals around the primary igneous minerals. A late stage of oxidative alteration is reflected in carbonate–haematite–smectite pseudomorphs of olivine and orthopyroxene.

Many of the gabbros have high palaeomagnetic intensities, particularly the magnetite–ilmenite varieties. The inclina-

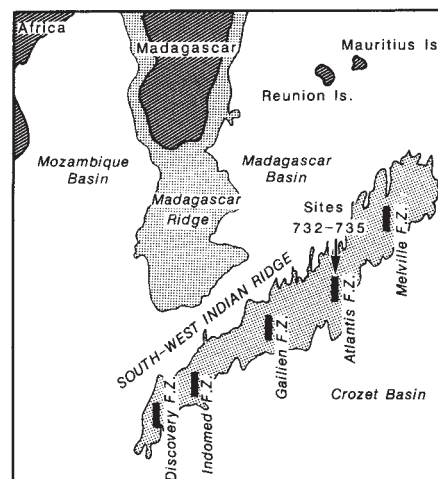


Fig. 1 Map showing land masses and ridges off the south-east coast of Africa and sites 732–735. F.Z., fracture zone.

tions are all steep and about equally divided between normal and reverse. Stable inclinations after demagnetization are reversed and average $65 \pm 15^\circ$, compared with a theoretical inclination of -52° for the site, based on an axially aligned dipole field.

Compressional-wave seismic velocities in the gabbros are in the range 6.5–7.0 km s⁻¹, both in rock samples and *in situ*, confirming that fracture porosity is not significant throughout the section. Directional anisotropy of up to 10 per cent occurs in the seismic properties, particularly in some of the foliated rocks. The electrical resistivity of the rock determined by logging in the hole varies from about 4 to $4 \times 10^4 \Omega$ m. The highest values are to be expected from silicate rock with low porosity, but the lower values are not generally correlated with higher porosity. The lowest resistivity is found in unit 4, and probably arises from the presence of iron–titanium oxides. Variations of about two orders of magnitude occur over distances of a few metres in deeper rocks.

A borehole packer experiment, in which part of the hole is isolated hydraulically and exposed to high fluid pressure, was used to determine fluid permeability. The section below about 272 m under the sea floor is relatively impermeable, but the permeability is higher by several orders of magnitude above. Because of the rock's low porosity, determined on

*Co-chief scientists: Paul T. Robinson (Dalhousie University) and Richard P. Von Herzen (Woods Hole Oceanographic Inst.); ODP Staff Scientist: Andrew C. Adamson (ODP, Texas A&M Univ.). Also Keir Becker (Univ. Miami), Sherman H. Bloomer (Duke Univ.), Mathilde Cannat (Univ. Bretagne Occidentale), Henry J.B. Dick (Woods Hole), Rolf F.K. Emmertmann (Inst. für Geowissenschaften und Lithosphärenforschung), Gunilla Gard (Univ. Stockholm), David Goldberg (Lamont-Doherty Geological Obs.), Rejean Hebert (Univ. Laval), Jan G.H. Hertogen (Univ. Leuven), Hartley Hoskins (Woods Hole), Gerardo Iturrino (Purdue Univ.), J. Dirk C. Kassenaar (Univ. Waterloo), Pamela D. Kempton (Open Univ.), Eiichi Kikawa (Univ. Tokyo), Stephen H. Kirby (US Geological Survey, Menlo Park), Peter S. Meyer (Woods Hole), James H. Natland (Scripps Inst. of Oceanography), Kazuhito Ozawa (Univ. Tokyo), Janet H. Pariso (Univ. Washington), James H. Scott (Lamont-Doherty), Debra S. Stakes (Univ. South Carolina) and Stephen A. Swift (Woods Hole).

both samples and from down-hole logging, the higher permeability is unlikely to be related to its bulk properties, but probably reflects fracture permeability at least in a few zones. The temperature gradient in the hole is very low or even negative, which could reflect the vertical oceanic temperature structure on this steep-sided

platform or slow seawater advection even in the low-porosity rocks.

Leg 118 has demonstrated the feasibility of drilling plutonic rocks with existing technology. There are many other favourable sites in fracture zones where sections of lower crust and upper mantle can be sampled. □

Nuclear physics

Laser-induced fission

J.E. Lynn

IN the 50 years since the discovery of fission (Hahn, O. & Strassmann, F. *Naturwissenschaften* 27, 11; 1939), powerful technologies based on the process have been developed, accompanied by detailed studies of the phenomenon. Many 'probes' for inducing nuclear fission have been found, the first being the neutron. Indeed, neutron-induced fission remains the basis of most uses of nuclear splitting. But other nuclear particles (protons, deuterons, α -particles and heavier nuclei) and non-nuclear beams such as γ -rays, electrons and muons are also effective. Now, according to Boyer, Luk and Rhodes (*Phys. Rev. Lett.* 60, 557–560; 1988), the possibility of inducing nuclear fission by ultra-violet light from an intense laser source should be added to this list. This would provide both a valuable tool for diagnosing the effects of intense fields on atomic electrons, and a bright point source of pulsed fission products and neutrons.

In principle, the possibility of inducing fission by light is not remarkable. The fission of nuclei heavier than iron is an exothermic reaction. Indeed, spontaneous fission is possible and is observed as a rare event for a few of the heaviest natural nuclei as well as for many artificial trans-uranic nuclei. But before splitting, a nucleus must first elongate from its equilibrium shape, an energy-expensive process creating a barrier against fission which even for uranium nuclei is 5–6 million electron-volts (MeV) high. Bombarding nuclear particles such as neutrons can provide most of this energy in the form of binding energy as they merge with the target nuclei, thereby creating excited compound nuclei with enough energy to overcome the fission barrier.

Non-nuclear beams provide the required energy directly to the target nucleus. Because such beams are of comparatively low intensity, the excitation process is by absorption of a single quantum of energy and, for electro-magnetic beams, this gives a requirement for radiation of very short wavelength. Suitable γ -ray beams are usually obtained by irradiating heavy-metal targets with electrons of several million electron-volts, but only a small fraction of the resulting *bremssstrahlung*

quanta (photons) are of sufficiently high energy to induce fission at appreciable rates. Electron-induced fission results from inelastic scattering of the electron by the Coulomb field of the nucleus, which excites the nucleus to the required energy. The required electron energies for this form of fission are ~ 10 MeV or more.

Muon-induced fission is a mechanism of particular interest. What chiefly distinguishes the muon from the electron is its mass (it is about 200 times heavier than the electron) as well as its short life (2 ms). With nuclei it forms muonic atoms, the muon replacing an orbiting electron. Because of the large mass of the muon, the muonic orbits are very tightly bound and transitions between the orbits have correspondingly higher energy than electronic orbital transitions. The innermost muonic orbitals of very heavy atoms penetrate the nucleus, and there is very strong electro-magnetic coupling between the nuclear and muon motion. Hence there is a high probability of a radiationless transition between the lowest muonic orbits sufficiently energetic to excite the nucleus to above its fission barrier.

The fission barrier of the muonic atom is greater than that of the bare nucleus by about 1 MeV owing to the change in muon binding energy following the increasing elongation of the nucleus that accompanies fission. The intra-orbital cascade following the absorption of a slowed muon into a muonic atom and the resulting fission occur very quickly. There is also a delayed fission process following the weak interaction of the muon with the nucleus, which transforms a proton into a neutron with the release of a large amount of excitation energy.

The process of optically induced fission, newly suggested by Boyer *et al.*, does not involve any similar form of direct electro-magnetic coupling to the nucleus. The postulated mechanism is an indirect one in which the intense focused ultraviolet light accelerates the atomic electrons in the target material to highly relativistic energies and electron-induced fission of the nucleus follows. The laser light thus acts as a miniature electron accelerator. It has been calculated (Feldman, M.J. & Chino,

R.Y. *Phys. Rev.* A4, 352; 1971) that a light intensity of 10^{12} W concentrated in a diffraction-limited focus would give an electric-field component of $\sim 3 \times 10^{11}$ V cm $^{-1}$ and this would raise an electron to relativistic energies in a single pass. At such energies the magnetic component of the optical field would bend the electron to move forward to travel with the wave.

Boyer *et al.* believe that the electro-fission resulting from a laser pulse of 1 J, 10^{-13} s duration focused in a spot of 1 μ m diameter on the surface of a solid uranium target will give a fission probability of 10^{-5} per nucleus. The penetration of ultra-violet light into the metal is relaxed by relativistic corrections but still limited to ~ 0.02 μ m. Thus about 8,000 fission events (equivalent to an energy yield of ~ 0.2 μ J) would occur in this spot.

As well as causing electrofission, the laser-accelerated electrons would produce *bremssstrahlung* electromagnetic radiation of sufficient energy to cause photofission. High-energy γ -rays penetrate about 1 cm into uranium, much further than light, and this would induce 10^6 fissions following photoabsorption, but only about $\sim 1,000$ of the fission products would escape from the front surface of the target. The fission-product yield is dominated by the electro-fission and the tiny volume in which it operates. The neutrons released by fission, on the other hand, are highly penetrating, thus the neutron yield is dominated by photofission, and the effective neutron source is spread over a large volume.

Although the report by Boyer *et al.* appears in the nuclear-physics section of *Physical Review Letters*, it seems unlikely that the main result of this phenomenon, if it is experimentally observed, will be to increase our understanding of the mechanism of fission, as the energy transferred by the scattered electron will be essentially unmeasurable. This is a pity because the spectroscopy of the mass and kinetic energy of the fission products should be highly measurable, given the small source size and short duration of the pulse, if suitable sub-picosecond time-of-flight techniques can be developed.

The relative ease and sensitivity of detection of fission events promise, however, to allow this phenomenon to be a valuable diagnostic technique for studying the coupling of intense laser light to atomic electrons. For applied physics, it gives the promise of an extremely small, pulsed source of fission-product nuclei and fast neutrons. Experimental realization may not be too far away, given that krypton-fluoride excimer lasers, yielding a wavelength of 248 nm, already operate at powers and pulse durations within an order of magnitude of those considered above (Schwarzenbach, A.P. *et al. Opt. Lett.* 11, 499; 1986). □

J.E. Lynn is in the Nuclear Physics Division, AERE Harwell, Didcot, OX11 0RA, UK.

Transfer RNAs

The second genetic code

Christian de Duve

A SECOND genetic code, still largely undeciphered, is written into the structure of aminoacyl-transfer RNA synthetases, enzymes that couple each amino acid to an appropriate transfer RNA molecule. The code is probably nondegenerate and could be older and more deterministic than the classical genetic code. On page 140 of this issue¹, Hou and Schimmel demonstrate one mechanism by which this second code can operate. These authors show that a single base pair can determine the amino-acid specificity of a transfer RNA molecule.

The accuracy of translation of genetic messages depends on the precision of two successive independent matchings: first, that of amino acids with transfer (t)RNAs; and second, that of charged tRNAs with ribosome-linked messenger RNAs. The latter is a relatively trivial example of direct interaction between two short nucleotide sequences, the anticodon and the codon, by base pairing. It is conducted entirely in the language of nucleic acids and is distinguished only by the fact that, because of

'wobbling', anticodons often recognize more than one codon. The matching between amino acids and tRNAs, on the other hand, is 'bilingual' and indirect. It is mediated by specific enzymes, aminoacyl-tRNA synthetases, which recognize the two partners and may therefore each be said to contain one entry of the genetic

that could be of fundamental importance. As more information becomes available about the recognition of tRNAs by aminoacyl-tRNA synthetases, it will be possible to test the validity of these speculations.

First, contrary to the classical code, the paracodon code is probably nondegenerate. There are only 20 aminoacyl-tRNA synthetases, each of which can recognize all the tRNAs (up to six) specific for a given amino acid. This recognition could involve different features in different tRNAs with the same amino-acid specificity, but this does not seem likely.

Second, consistent with its nondegeneracy, the paracodon code could be more deterministic than the classical code; it could even depend on stereochemical interactions between the paracodons and the cognate amino acids. A particularly simple kind of interaction would be one between the paracodon and an enzyme-bound aminoacyl-adenylate or an aminoacyl-enzyme intermediate. The fact that aminoacyl-adenylate formation precedes tRNA recognition² is consistent with this possibility. So is the fact that tRNA is involved in editing out

mismatched aminoacyl groups³. It is also possible that the region of interaction in the synthetase could include a crucial residue of the appropriate nature. Even if more complex recognition patterns are involved, the possibility remains that paracodons may have evolved from stereochemically determined 'proto-paracodons'.

Third, in agreement with these proposals, the paracodon code could be older than the classical code. Several authors have speculated⁴⁻¹¹ that tRNAs arose as short aminoacyl-carrying oligonucleotides that owed their specificity to direct interactions with the amino acids. Instead of being lost in evolution, the structural features involved in these interactions could have survived in the paracodons. The anticodons would have appeared

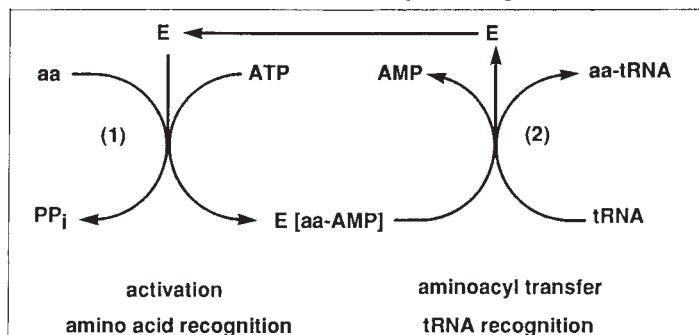


Fig. 1 The two steps of the reaction catalysed by aminoacyl-tRNA synthetases. The tRNA is recognized in the second step (through the features designated paracodon in this article) by the enzyme carrying a bound aminoacyl-AMP intermediate. Participation of the aminoacyl group in the recognition process is thus an attractive possibility.

dictionary.

The structural features of tRNAs recognized by aminoacyl-tRNA synthetases are still poorly characterized. But an important conclusion already emerging (refs 2, 3) is that in several instances, the recognition site in the tRNA does not include the anticodon. The paper by Hou and Schimmel in this issue demonstrates this point in a particularly clear-cut fashion, by showing that alanine specificity depends critically on a single base pair, G3U70, in the acceptor stem of the *Escherichia coli* tRNA^{ala} molecule, and that this specificity can even be conferred on two other tRNAs by introduction of this base pair into their molecules. Normanly *et al.*⁴ showed two years ago that a tRNA^{leu} could be converted to a tRNA^{ser} by 12 nucleotide replacements not involving the anticodon. They did not establish whether all replacements (which included the 3-70 pair) were critical.

These results point to the existence of a second genetic code, imprinted into the structure of aminoacyl-tRNA synthetases and matching the amino acids with the structural features of tRNAs that are recognized by these enzymes (see figures). For the purpose of discussion, I shall call these features paracodons. The choice of this term is not meant to imply that paracodons are nucleotide triplets similar to codons or anticodons, but the term is convenient for defining a code that is still waiting to be deciphered.

This second code has several features

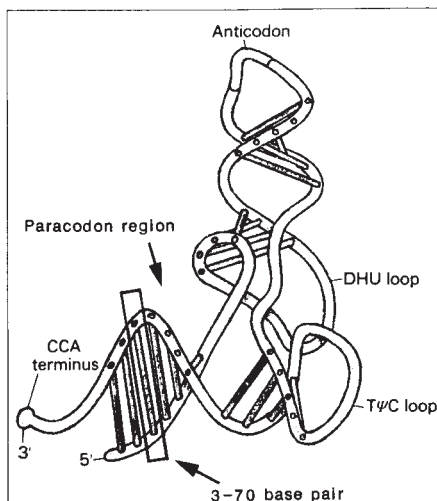


Fig. 2 Schematic representation of the three-dimensional conformation of a tRNA. The paracodon region designates the general area where sites of recognition by aminoacyl-tRNA synthetases have been tentatively located^{2,3}. The box indicates the position of the G3U70 base pair found to be crucially important by Hou and Schimmel¹. Note that the anticodon is about as far removed from the aminoacyl-accepting CCA 3'-terminus as is possible. In contrast, a direct interaction between paracodon and aminoacyl group is sterically conceivable. From ref. 13.

- Hou, Y.-M. & Schimmel, P. *Nature* **333**, 140-145 (1988).
- Schimmel, P. & Söll, D. *A. Rev. Biochem.* **48**, 601-648 (1979).
- Schimmel, P. *A. Rev. Biochem.* **56**, 125-158 (1987).
- Normanly, J., Ogden, R.C., Horvath, S.J. & Abelson, J. *Nature* **321**, 213-219 (1986).
- Woese, C. *The Genetic Code* (Harper & Row, New York, 1967).
- Woese, C. *J. molec. Evol.* **3**, 109-113 (1974).
- Dillon, L.S. *The Genetic Mechanism and the Origin of Life* (Plenum, New York, 1978).
- Ninio, J. *Approches Moléculaires de l'Évolution* (Masson, Paris, 1979).
- Folsome, C.E. *The Origin of Life* (Freeman, San Francisco, 1979).
- Dyson, F. *Origins of Life* (Cambridge University Press, 1985).
- de Duve, C. in *The Roots of Modern Biochemistry*, Lipmann Memorial Symp. (eds Kleinkauf, H. von Döhren, H. & Jaenicke, L.) 881-894 (de Gruyter, Berlin, 1988).
- Crick, F.H.C. *J. molec. Biol.* **38**, 367-379 (1968).
- Quigley, G.J. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 4866-4870 (1975).

later, during the further development of the tRNA molecules. This eventuality has interesting implications, including the possibility of early polypeptide replication and even reverse translation¹¹. In any case, the development of the classical code as a "frozen accident" (ref. 12) becomes more easily understandable if it was superimposed on a more fundamental, deterministic code.

Finally, the two codes are linked by the physical association of anticodons and paracodons in the same tRNA molecules. □

Christian de Duve is Director of the International Institute of Cellular and Molecular Pathology in Brussels and Andrew W. Mellon Professor at the Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA.

Fluid dynamics

Meteorology of a flat Earth

Ian N. James

IN a recent paper¹, D. L. Boyer and R.-R. Chen claim that several features of the climate flow in the Northern Hemisphere can be reproduced in a laboratory apparatus. Their experiment involves towing models of the Rockies, Himalayas and Greenland through a rotating tank containing stratified salt solution to represent the troposphere. Such an arrangement differs from that of the real Earth in several respects, most important of which are the absence of spherical geometry and of temperature gradients. The experiments, however, successfully reproduce the main troughs and ridges of pressure found in association with the Rockies and Himalayas (see figure), and a characteristic oscillation with a period that corresponds to one of about 40 days which has been identified in atmospheric data².

The experiments reported were carefully designed so that 15 dimensionless parameters take similar values to those characterizing atmospheric circulation. These include measures of rotation rate and stratification. Of course other parameters such as those measuring the typical horizontal temperature gradients and the curvature of the Earth could not take atmospheric values. Both these, but especially the latter, are generally accepted as being crucial in accounting for the observed patterns of troughs and ridges in planetary-scale flow. The most successful theories of the longitudinal variations of the mean flow are in terms of steady trains of 'Rossby' waves excited by mountain ranges and land-sea contrasts³.

The most remarkable results of the experiments are the successful reproduction of the troughs and ridges associated with the Rockies and Himalayas, and the characteristic oscillation. This oscillation period is close to the time for elements of fluid to circumnavigate the apparatus; its persistence implies that dynamical processes continually replenish the eddy energy of the travelling disturbances.

A rotating fluid supports various kinds of wave motion, characterized by different frequencies. The kind of waves excited by a mountain depend on the typical

frequency with which it deflects air parcels. In the atmosphere, it takes a day or two for air to pass over mountain ranges such as the Rockies, with the result that mainly low-frequency 'Rossby' waves (north-south transverse waves) are forced. Internal gravity (or buoyancy) waves have much shorter periods, in the range 10–20 minutes, and so are not significantly excited. Air flow over much smaller hills (such as individual ridges and valleys within the Rockies) is on this sort of timescale and there has recently been renewed



Horizontal streak pattern from the dish-pan experiment of Boyer and Chen. The solid white objects are the forms representing the Rockies, the Himalayas and Greenland. (From ref. 1.)

interest in the gravity waves forced by the roughness of the main mountain ranges⁴. But in the laboratory experiments, Rossby waves could not exist because they require curvature — absent from the system — to propagate. Without any substantial wave propagation, disturbances would be confined to the region close to the artificial mountains; it must be that other kinds of wave motion generate the distant effects observed. The question remains whether Boyer and Chen obtained their apparently realistic results through the operation of very different dynamics. The radiation of stationary internal gravity waves would seem to be a possible mechanism.

The reproduction of a 40-day oscillation is interesting. Current theories of the

oscillation suggest it is an essentially tropical phenomenon, driven by the interaction of Kelvin waves (buoyancy-driven edge waves trapped on the equator) with latent heating in convective cloud systems. The problem with this theory is that most computer models designed to test it produce an oscillation with much too short a period. Triggering by disturbances in the mid-latitude flow, with an essentially advective timescale, could be important in increasing the period to that observed.

The paper is thought-provoking; the issues it raises are perhaps rather philosophical. In particular, I worry about the validity of laboratory analogues of geophysical flows. That Chen and Boyer's experiments leave out certain important effects present in the atmosphere does not matter — it could help to separate important from secondary effects. But diagnosis of the flow must be subtle enough to distinguish dynamically significant parallels from merely superficial resemblance. I question whether they have succeeded.

More fundamentally, ensuring dynamical similarity may not be enough. The boundary conditions and geometry of the apparatus could also be important. This is certainly true in the case of the radiation of Rossby waves from mountains. The reflection of such waves from the rigid walls of a laboratory apparatus contrasts with the behaviour of such waves as they propagate into the Earth's tropics. As the westerly winds of the mid-latitudes are replaced by the easterlies typical of the tropics, a 'critical line' where the local flow speed matches the phase speed of the Rossby waves is encountered. The behaviour of the wave train near the critical line is complex and poorly understood. If enough dissipation is present, it will be absorbed. In inviscid flow, a highly nonlinear, amplitude-dependent partial reflection may occur⁵.

Laboratory studies have an honourable history in geophysical fluid dynamics. But one can question whether they continue to be a fruitful line of enquiry. Only when geometrical as well as dynamical similarity can be guaranteed will they provide exact analogues. Technical and practical difficulties rarely allow this. On the other hand, the more abundant geophysical data now available, interpreted in conjunction with mathematical and numerical models of varying complexities, do enable the fluid dynamics of natural geophysical fluid systems to be explored⁶. □

1. Boyer, D.L. & Chen, R.-R. *J. Atmos. Sci.* **44**, 3552 (1987).
2. Madden, R.A. & Julian, P.R. *J. Atmos. Sci.* **29**, 1109 (1972).
3. Held, I.M. in *Large-scale dynamical processes in the atmosphere* (eds Hoskins, B.J. & Pearce, R.P.) 127–168 (Academic, London, 1983).
4. McFarlane, N.A. *J. Atmos. Sci.* **44**, 1775–1800 (1987).
5. Killworth, P.D. & McIntyre, M.E. *J. Fluid Mech.* **161**, 449–492 (1985).
6. Hoskins, B.J. *Q. J. R. Met. Soc.* **109**, 1–22 (1983).

Ian N. James is in the Department of Meteorology, University of Reading, PO Box 239, Reading RG6 2AU, UK

Geology

Evolution of the lower crust

Stephen M. Wickham

GRANULITES, high-grade metamorphic rocks that are deficient in hydrous minerals, are significant constituents of the middle and lower continental crust, and are of great petrological interest¹⁻⁴. Elucidating the conditions in which they formed will reveal much about the genesis of the early crust and its subsequent evolution through episodes of melting and metamorphism. Current research has been provoked largely by the development of models⁵ of granulite petrogenesis involving flushing the lower crust with large volumes of mantle-derived CO₂. Jackson *et al.*, on page 167 of this issue⁶, report new evidence that supports this model. They indicate a two- to fourfold increase in CO₂ abundance for fluid inclusions in quartz from charnockite veins and patches within amphibolite gneiss, and marked changes in carbon isotope compositions.

This and other studies were the basis of a recent workshop* in India, which was organized around extensive field trips to exemplary outcrops. The rocks that were examined support a wide range of models of granulite formation.

The CO₂-flushing model⁵ of granulite petrogenesis is based primarily on the observation⁷ that many granulites contain CO₂-rich, high-density fluid inclusions, and on the recognition in southern India of transition zones in which amphibolite facies gneisses (high-grade rocks containing hydrated silicates) are partially dehydrated and almost isochemically transformed into CO₂-inclusion-rich charnockites (granulites containing the anhydrous mineral orthopyroxene). This model has important ramifications for many branches of Earth sciences, but it is highly controversial.

Participants at the workshop identified textural evidence for several mechanisms of granulite formation (CO₂ infiltration, dehydration melting and localized exsolution of CO₂ from silicic magma) during a 14-day traverse from Bangalore to Trivandrum. At some sites, features could be explained only by a combination of these processes. Important new evidence was found at several localities indicating the generation of granulite assemblages by dehydration melting (albeit on a local scale). But stable-isotope data from the

same area, presented by Jackson *et al.* in this issue⁶, support the involvement of a CO₂-rich fluid.

South of the amphibolite-granulite transition zone, in the main granulite terrane, massive charnockites outcrop over wide areas. Some participants doubted that these rocks represent transformed (CO₂-fluxed) pre-existing crust, and preferred to interpret them as large CO₂-rich plutons. CO₂ released from such bodies on crystallization could have given rise to the patchy charnockitization

At the same localities, there is also much evidence of conversion of amphibolite to charnockite with no textural evidence of melting. Such charnockite patches, having a much higher CO₂ abundance and higher $\delta^{13}\text{C}$ values⁶ (the relative deviation of the ¹³C/¹²C isotope ratio from the 'PDB' standard) than adjacent gneisses, probably result from an influx of externally derived CO₂-rich fluid. The consistent orientation of charnockite patches and streaks at some outcrops implies that the influx was fracture controlled. It is possible that fluid infiltration itself could also trigger local dehydration melting, adding a yet further complication to the interpretation of textures.

The partially charnockitized localities are consistent with granulite petrogenesis

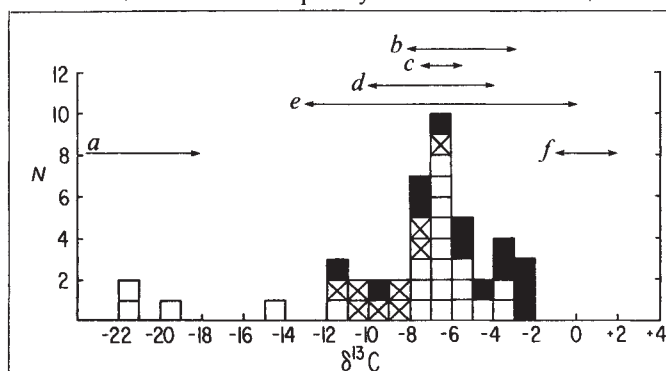


Fig. 1 Histogram of the carbon isotope composition of CO₂-rich fluid inclusions from granulite terranes. The $\delta^{13}\text{C}$ values (the relative deviation of the ¹³C/¹²C ratio from the 'PDB' standard) typically fall in the range -2 to -12, overlapping those of the mantle, continental hydrothermal sources and metamorphosed crustal carbonate. Typical values for organic carbon (a), continental CO₂-emanations (b), mid-ocean-ridge basalts (c), carbonates and diamonds (d), Naxos fluid inclusions¹⁶ (e) and marine carbonates (f) also shown. Clear squares, from refs 14 and 15; crossed, ref. 6; black, ref. 2.

observed at higher structural levels.

At the partially charnockitized localities, spots and streaks of grey charnockite occur within amphibolite-grade gneisses, but typically constitute less than half of the total outcrop area. The textural evidence of local melting at some of these localities, with concomitant development of orthopyroxene, is compelling but is typically developed on a small (centimetre) scale. For melting to generate large volumes of granulite, the water originally bound in hydrous minerals must first dissolve in the melt and then escape from the rock when the melt crystallizes without 'back-reacting' to reconvert charnockite to amphibolite. This is unlikely at granite melting temperatures. Alternatively, the melt itself must be extracted from the rock; but compaction theory⁸ predicts that, at low melt-fraction, silicic liquids (particularly high-viscosity, water-undersaturated ones) separate only slowly from residual solid material and that they are incapable of stripping kilometric regions of their melt over realistic time-scales.

on a local scale only. Although there is a gradual transition southwards to regionally extensive massive charnockites, it is a leap of faith to interpret the latter lithologies as equivalent protoliths that have been completely converted. Data suggest that the massive charnockites are indeed rich in CO₂, mostly present in fluid inclusions. In the Nilgiri Hills, some granulites contain several volume per cent of fluid-inclusion CO₂ (ref. 9) and abundant, tiny, late-stage calcite veinlets in brittle fractures. Such veinlets are common in many granulite terranes (J. L. R. Touré, personal communication) and could represent CO₂ liberated from burst fluid inclusions during uplift. This raises the intriguing possibility that estimates of CO₂ contents of such rocks based on fluid inclusions are low, because of losses during subsequent uplift.

Whether the Indian massive charnockites are transformed (carbonically metamorphosed) gneisses or CO₂-rich plutons, CO₂-rich fluids clearly were intimately involved in their petrogenesis. CO₂ contents as high as those in the Nilgiri Hills represent massive enrichment, and if present in large volumes of the lower crust, demand an explanation for their source. Experiments on the solubility of mixed H₂O-CO₂ fluids in silicate melts will reveal whether all the CO₂ could originally have been dissolved in charnockitic magmas, or whether it would have had to stream through CO₂-saturated charnockitic magmas. Emplacement of these magmas into the crust, followed by degassing during crystallization, could have left abundant CO₂-rich inclusions in the massive charnockites, and also caused the patchy charnockitization at higher levels.

The $\delta^{13}\text{C}$ values of granulite CO₂-rich fluid inclusions overlap those of mantle

* Deep continental crust of south India, Bangalore-Trivandrum, 9-23 January 1988.

carbon (Fig. 1). A normal mantle source, however, seems insufficiently rich in carbon to explain the observed CO_2 enrichment of granulites. Nor could it generate the CO_2 flux necessary to convert amphibolite to granulite¹⁰. Furthermore, metamorphosed crustal carbonates also have $\delta^{13}\text{C}$ values in the same range¹⁰. Although marine limestone has $\delta^{13}\text{C} \approx 0$, this value will be driven down to negative values by partial decarbonation caused by regional or contact metamorphism, or by exchange with organic carbon or juvenile CO_2 . High-grade metacarbonates with a long crustal history (as might be expected for those in the deep crust) typically have

granitic magma bodies, and episodic CO_2 fluxes, consistent with the inferred petrogenesis of southern Indian charnockites. This model is particularly appropriate if some of the massive charnockites are indeed magmas with high CO_2 contents (either dissolved or present as a free vapour phase) rather than transformed metasediments. CO_2 -rich magmas and associated large CO_2 fluxes could be a natural consequence of crustal growth processes involving underplating of mantle material beneath a carbonate-bearing lower crust.

It is interesting to note that large CO_2 fluxes are observed in hydrothermal

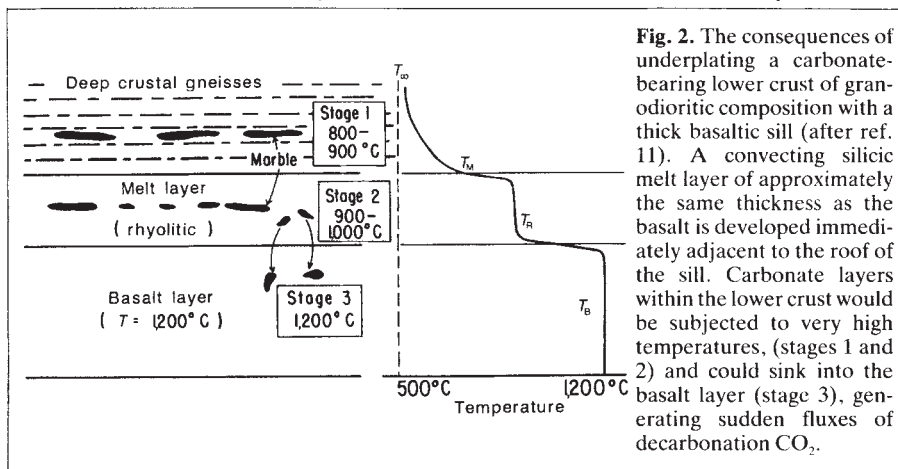


Fig. 2. The consequences of underplating a carbonate-bearing lower crust of granodioritic composition with a thick basaltic sill (after ref. 11). A convecting silicic melt layer of approximately the same thickness as the basalt is developed immediately adjacent to the roof of the sill. Carbonate layers within the lower crust would be subjected to very high temperatures, (stages 1 and 2) and could sink into the basalt layer (stage 3), generating sudden fluxes of decarbonation CO_2 .

$\delta^{13}\text{C}$ values in the range +2 to -10. Carbon isotope data, therefore, do not distinguish a unique source for the CO_2 , although it must be exotic to the sample locality⁹.

Further systematic analysis of carbon isotopes in CO_2 in both fluid inclusions^{2,4,6} and cordierite¹ from localities at various structural levels throughout the terrane, combined with regional oxygen- and strontium-isotope studies, could help to constrain its origin further.

At least two and possibly three different charnockite-forming events, each involving a CO_2 influx at least on a local scale, have affected the deep crust of southern India. This suggests that the CO_2 source lies in the Indian crust or mantle lithosphere. Either could be enriched in carbonate; Central Highland Series granulites in Sri Lanka contain 5 per cent carbonate layers by volume, and mantle-derived carbonatites and other alkaline igneous rocks occur sporadically throughout southern India. A crustal carbonate source is probably quantitatively more significant and could be mobilized during underplating and partial melting of the lower crust (Fig. 2). Underplating could generate large-scale melting¹¹ at 900–1,200 °C and promote high-temperature, total decarbonation of partially decarbonated marble layers (stages 1, 2 and 3 in Fig. 2). The combined process would result in high-grade metamorphism, melting, emplacement of CO_2 -saturated

systems associated with zones of active extension and volcanic activity, such as the Massif Central¹², where mafic intrusion and underplating might be expected to be occurring at depth. $^{13}\text{C}/^{12}\text{C}$ isotope ratios are typically in the same range as that shown by granulite fluid inclusions, but no significant ^3He anomaly is observed, implying that the source for the emanations is dominantly crustal. This further emphasizes the importance¹³ of metacarbonate layers in the deep continental crust as a significant source for large fluxes of CO_2 . □

1. Vry, J., Brown, P.E., Valley, J.W. & Morrison, J. *Nature* **332**, 66–68 (1988).
2. Santosh, M., Hoefs, J. & Iyer, S.S. *Nature* (submitted).
3. Frost, B.R. & Frost C.D. *Nature* **327**, 503–506 (1987).
4. Stahle, H.J., Raith, M., Hoernes, S. & Delphs, A. *J. Petrol.* **28**, 803–834 (1987).
5. Newton, R.C., Smith, J.V. & Windley, B.F. *Nature* **288**, 45–50 (1980).
6. Jackson, D.H., Matthey, D.P. & Harris, N.B.W. *Nature* **333**, 167–170 (1988).
7. Touret, J.L.R. in *The Deep Proterozoic Crust of the North Atlantic Provinces* (eds Tobin, A.C. & Touret, J.L.R.) 517–549 (NATO ASI Series C, **158**, 1985).
8. Wickham, S.M. *J. geol. Soc. Lond.* **144**, 281–297 (1987).
9. Srikanthappa, C. *J. geol. Soc. India* **30**, 72–76 (1987).
10. Wickham, S.M. *J. geol. Soc. India* **31**, 162 (1988).
11. Huppert, H.E. & Sparks, R.S.J. *J. Petrol.* (in the press).
12. Matthews, A., Fouillac, C., Hill, R., O'Nions, R.K. & Oxburgh, E.R. *Earth planet. Sci. Lett.* **85**, 117–128 (1987).
13. Glassley, W.E. *Contrib. Mineral. Petrol.* **84**, 15–24 (1983).
14. Hoefs, J. & Touret, J.L.R. *Contrib. Mineral. Petrol.* **52**, 165–174 (1975).
15. Hoefs, J., Coolen, J.J.M. & Touret, J.L.R. *Contrib. Mineral. Petrol.* **78**, 332–336 (1981).
16. Kreulen, R. *Am. J. Sci.* **280**, 745–771 (1980).

Stephen M. Wickham is in the Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA.

Daedalus

Fine spread of colour

PAINTING is one of the most expensive, tedious and unsatisfactory of technical operations, whether conducted by large-scale dip coating and stoving, or by slow old brush. Worse still, the paint is bound to fail or wear off sooner or later, when it must all be done again.

Daedalus, however, feels that painting ought to be easy. A liquid must spread spontaneously over a surface if the process releases surface energy. Some liquids spread on other liquids in this way; but surface energies are generally too low (less than 1 joule per square metre) to drag a liquid over a solid surface at a useful rate, even if the energy balance favours it.

But, says Daedalus, suppose the liquid reacts with the solid surface. Chemical energies are so great that spreading could now be favoured by many kilojoules of released energy per square metre. A chemically reactive paint should not only spread strongly across a horizontal surface: it should even be able to haul itself up a vertical surface against gravity.

To test this cunning notion, Daedalus is devising a self-spreading paint specifically for metal surfaces. DREADCO's chemists are compounding polysulphide paints which react with a metal surface, binding the paint firmly with sulphide bonds and releasing much energy in the process. The resulting 'autopaint' should spread vigorously over metal and adhere with great tenacity. A metal object dipped even partially into autopaint will soon be completely covered by the rapidly spreading film.

Autopaint will need careful handling. It will tend to climb out of its can and spread relentlessly over all available metal. A region to be coated with autopaint will need a boundary drawn round it in wax crayon, as a barrier to further progress. Even so, autopaint with its chemical adhesion may simply burrow under and displace the merely physically adhering wax, and continue ruthlessly on its way. But this alarming characteristic may also be a virtue. Applied to a dirty or rusty surface, autopaint will burrow under the grime and protect the metal perfectly from beneath.

Thus autopaint will rescue all existing conventionally painted metal objects: lamp-posts, locomotives, gas-holders, bridges. Simply expose the bare metal in a few places, and keep them topped up with autopaint as it spreads all over the surface under the existing paint. The structure will then be perfectly preserved from the elements, but will still look just as tatty. Industrialists will be content; but car owners, for whom appearance is usually more important than mechanical durability, will probably still prefer old-fashioned repainting methods.

David Jones

Interlayer effects in high- T_c superconductors

SIR—A striking feature of the known high-transition-temperature superconductors is the presence of square-planar Cu–O layers. In the newly discovered thallium¹ and bismuth² families, one, two or three of these Cu–O layers are sandwiched between layers of thallium or bismuth oxide³ and the superconducting transition temperature (T_c) increases with the number of layers in the sandwich. This points to an interlayer mechanism such as the one we have recently discussed⁴.

The basic ingredient in our model is the coherent hopping of valence-bond pairs⁵ between Cu–O planes. T_c is set principally by the values of the parameters $\Lambda = t_1^2/J$, where t_1 is the interlayer hopping matrix element and J is the intralayer exchange constant. It is believed that phase coherence between valence bonds can be achieved by pairing the holons, which are bose-like in the background of resonating valence bonds. In this way it is hoped that the essential physics of the isolated layer (strong correlations and pre-formed pairs) can be handled correctly both above and below T_c . Of course it is not clear *a priori* whether there might exist a stronger purely in-plane pairing effect; one must appeal to experiment to settle this issue.

Let n be the number of Cu–O sublayers. The superconducting amplitude $F_i = \Sigma_i \langle e_i e_{-i-k} \rangle$ (where e is the holon destruction operator) must be the same in all layers (l) for $n = 1$ and $n = 2$, as in these cases all layers are equivalent. For $n = 1$ and $n = 2$, we find $T_c = 2\Lambda/p\delta$ and $T_c = (\Lambda_o + \Lambda_m)p\delta$, respectively, where p is a model-dependent number of order unity and δ is the doping away from nominal valence Cu^{2+} . The couplings are defined in the figure, and their values derived from the observed T_c are given in the table legend.

In contrast, when $n = 3$ or 4 there are inequivalent outer (o) and middle (m)

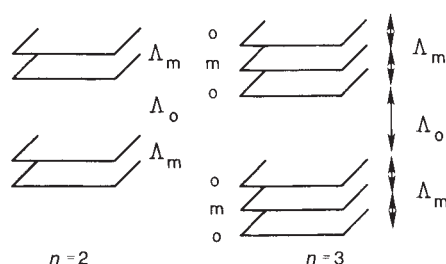
sublayers; nevertheless, all layers must go superconducting at a single temperature. T_c is given in these cases by the eigenvalue equations

$$\frac{T_c}{p\delta} \begin{pmatrix} F_m \\ F_o \end{pmatrix} = \begin{pmatrix} 0 & 2\Lambda_m \\ \Lambda_m & \Lambda_o \end{pmatrix} \begin{pmatrix} F_m \\ F_o \end{pmatrix} \quad n=3$$

$$\frac{T_c}{p\delta} \begin{pmatrix} F_m \\ F_o \end{pmatrix} = \begin{pmatrix} \Lambda_m & \Lambda_m \\ \Lambda_m & \Lambda_o \end{pmatrix} \begin{pmatrix} F_m \\ F_o \end{pmatrix} \quad n=4$$

The predicted T_c s for the $n = 3$ bismuth and thallium materials, shown in the table, are in good agreement with experimental values.

At $T = 0$ another clear test of the interlayer mechanism emerges. It is easily shown that the superconducting amplitudes F are equal on all layers at $T = 0$.



Cu–O sublayers and definition of the pair-hopping parameters, Λ_o and Λ_m .

This implies distinct order parameters $\Delta_m/\Delta_o = 2\Lambda_m/(\Lambda_m + \Lambda_o)$ for inequivalent planes when $n \geq 3$. $\Lambda \approx \Lambda F$ determines the change spectrum in superconducting state. A purely in-plane mechanism necessarily leads to a single Δ .

Finally we comment on the couplings found. The Λ_m s for the bismuth and thallium compounds are similar, as is expected from their structure. Λ_o for Bi is small compared with that of Tl, perhaps because the Tl p -band is closer to Cu levels than is the case for Bi. (There is also the cleavage difference, as noted in ref. 3.) Note, however, that the simple numerology above assumes that the doping is the same in all layers, and that the doping is roughly the same for the Bi and Tl compounds for the quoted T_c s.

J.M. WHEATLEY

T.C. HSU

P.W. ANDERSON

Joseph Henry Laboratories of Physics,
Princeton University,
Princeton, New Jersey 08540, USA

- Sheng, Z.Z. & Hermann, A.M. *Nature* **332**, 58–59; 138–139 (1988).
- Maeda, H., Tanaka, Y., Fukutomi, M. & Asano, T. *Jap. J. appl. Phys. Lett.* (submitted).
- Hazen, R.L. *et al. Phys. Rev. Lett.* **60**, 1657–1660 (1988).
- Wheatley, J.M., Hsu, T.C. & Anderson, P.W. *Phys. Rev. B* **37**, 5897–5900 (1988).
- Anderson, P.W. *Science* **235**, 1196–1198 (1987).
- Poole, R. *Science* **240**, 146–147 (1988).

Observed ($n=1,2$) and predicted ($n=3,4$) transition temperatures for bismuth and thallium superconductors

n	Formula*	T_c (K)	
		Bi	Tl
1	2021	10	80
2	2122	85	105
3	2223	116	114
4	2324	131	119

The observed T_c s for the 2012 and 2122 compounds (ref. 6 and J.B. Torrance *et al.*, personal communication) yield $\delta \Lambda_o = 5$ K for Bi, 40 K for Tl; $p\delta \Lambda_m = 80$ K for Bi, 65 K for Tl. The predicted T_c s for the 2223 and (hypothetical) 2323 compounds are then calculated from these parameters.

* The four digits represent the number of cations in the unit cell, in the order Bi(Tl), Ca, Sr(Ba), Cu.

Solar cycle 22 to be one of the largest on record?

SIR—It now seems virtually certain that the value of 12.3 for the smoothed monthly mean sunspot number for September 1986 defines the minimum of solar cycle number 21, and hence the conventional beginning of cycle 22. Since then, activity has increased unusually steeply. Because it has generally been found that rapid initial rises in sunspot activity have been followed by large maxima, there is already wide interest in the development of cycle 22.

Given the comparative regularity of the solar cycle period, the prediction of the timing of a subsequent sunspot maximum does not pose such a problem as the prediction of its magnitude. The consensus seems to be that cycle 22 will peak in 1990 ± 1 . Magnitude predictions are based on a variety of techniques¹, which generally give a large spread of values. Recent predictions^{2–5} for cycle 22 have all been at the top end of the range (International Sunspot Number as great as 190), compared with tentative earlier anticipations of a very small maximum (even as low as 36).

One prediction technique which has a broad scientific basis, and which generally gave near-correct forecasts for the previous cycle¹, makes use of experimentally observed 'precursors'. According to the dynamo theory, a given solar cycle is generated in the Sun by the gradual conversion of a poloidal magnetic field into one of toroidal form, powered by the Sun's differential rotation during most of the declining phase of the preceding cycle. If this is so, then solar or geophysical data obtained during this phase might be expected to contain clues about the forthcoming activity. One such precursor method^{6–8} is based on the occurrence of so-called geomagnetic abnormal quiet days (AQD)⁹ determined from the form of the diurnal variation of the magnetic horizontal component on international quiet days at a temperate latitude station, such as Hartland (51° N, 4° W geographic).

The annual number of AQDs (n) varies inversely with the solar cycle, and the peak number that occurs near a sunspot minimum anticipates the size of the subsequent sunspot maximum. This relationship holds good back to 1885, the earliest year for which data are available⁷. To illustrate the relationship, Fig. 1 shows an example of the variation of n over the last few solar cycles. Due to the necessary smoothing of the data, n is always around two years in arrears, so that it is only possible to locate the maximum with confidence some three years after it occurred. It is clear that the peak value of n that anticipates cycle 22 may now be identified fairly confidently. Fortunately for prediction purposes, this maximum has occurred

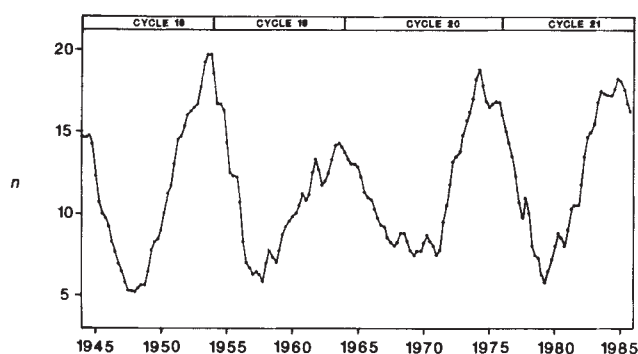


Fig. 1 Variation of annual number of AODs (n) in solar cycles 18–21.

before the end of the cycle, as in several previous cycles.

Prediction is based on the good linear relationship between the annual mean maximum sunspot number of a cycle Ra_{max} and the increase (Δn) in n over the declining phase of the preceding cycle. A regression line based on cycles 13–20 can be found in ref. 10. Subsequent analysis for cycle 21 shows that the method led to an accurate prediction for this cycle. Assuming an uncertainty of $\pm 20\%$ in estimating Δn , the peak sunspot number of each cycle has been successfully mirrored by this parameter determined near the preceding sunspot minimum. The value of Δn available from Fig. 1 as the precursor to cycle 22 leads to the prediction that for cycle 22 $Ra_{max} \approx 174 \pm 35$.

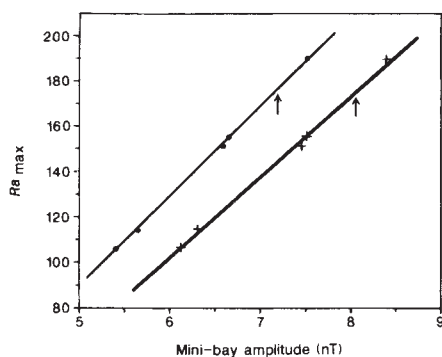


Fig. 2 Linear relation between average mini-bay amplitude near sunspot minimum and the magnitude of the following annual mean sunspot maximum. Amplitudes measured from a base level given by the average of the five values on either side of zero hour (circles) or from a base level given by the average of the four values on either side of ± 1 h (crosses). Arrows, corresponding amplitudes which anticipate cycle 22.

Erratum

IN the letter "Evidence for global warming in the past decade" by P. D. Jones *et al.* (*Nature* **332**, 790; 1988), part of a sentence in the second paragraph was inadvertently omitted. The correct version should read "Increases of CO_2 and other radiatively active trace gases . . . are expected to raise global mean temperatures. Increases of between 1.5 and 4.5 °C are expected from equilibrium general circulation models¹ for doubling of atmospheric CO_2 ." □

It has been established¹¹ that one of the main causes of an AOD is the imposition of a short-lived southward magnetic field, somewhat like that associated with a substorm, but of much smaller magnitude. My collaborators and I refer to the phenomenon as a 'mini-bay' because it resembles a long-duration negative bay of very small amplitude (typically 5–6 nT only).

We have previously shown⁸ that the main factor in explaining the ability of AOD occurrence to predict solar-cycle amplitudes is that the average magnitude of mini-bay amplitude on AODs near a sunspot minimum relates to the size of the next sunspot maximum. On this basis, the AOD count transcribes into an index measuring the extent of mini-bay activity, and hence the average AOD mini-bay amplitude joins the AOD count as a precursor, consistent with the dynamo theory of the production of a solar cycle.

Figure 2 shows the relation between the average mini-bay amplitude on AODs, determined for the years around each sunspot minimum for the last five cycles, and the subsequent value of Ra_{max} . The two lines use data based on different criteria for the choice of zero for the amplitude scale, and both are remarkably linear. The vertical arrows show the mini-bay amplitudes estimated for the current solar minimum epoch (1985–87) using the same two zero criteria, and it is evident that the prediction does not depend critically on the base level selected for the measurement of mini-bay amplitude, since both give identical predictions of $Ra_{max} \approx 175 \pm 35$. It seems likely that cycle 22 could be second only to cycle 19 as the largest cycle on record.

GEOFFREY M. BROWN

Department of Physics,
University College of Wales,
Aberystwyth SY23 3BZ, UK

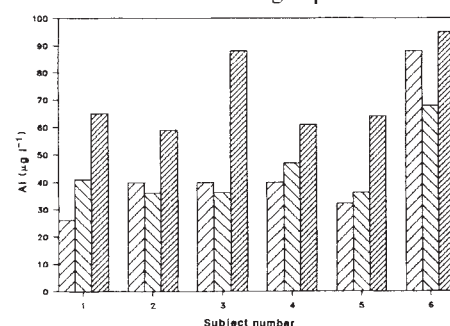
1. Brown, G.M. *Proc. Workshop on Solar-Terrestrial Predictions Meudon 1984* 1–7 (NOAA, Boulder, 1986).
2. Thompson, R. J. *Ionospheric Predictions Service Report IPS-TR-85-06* (Sydney, 1985).
3. Schatten, K.H. & Sofia, S. *Geophys. Res. Lett.* **14**, 632–635 (1987).
4. Kane, R.P. *Solar Phys.* **108**, 415–416 (1987).
5. *Solar Geophysical Data No. 516*, part 1, 14–17 (NOAA, Boulder, 1987).
6. Brown, G.M. & Williams, W.R. *Planet. Space Sci.* **17**, 455–470 (1969).
7. Brown, G.M. *Nature* **251**, 592–594 (1974).
8. Brown, G.M. & Butcher, E.C. *Planet. Space Sci.* **29**, 73–77 (1981).
9. Brown, G.M. *Proc. Workshop on Solar-Terrestrial Predictions Meudon 1984* 41–47 (NOAA, Boulder, 1986).
10. Brown, G.M. *Proc. Workshop on Solar-Terrestrial Predictions Boulder 1979* Vol. II, 264–279 (NOAA, Boulder, 1979).
11. Butcher, E.C. & Brown, G.M. *Geophys. J. R. astr. Soc.* **64**, 513–526 (1981).

Increased urinary excretion of Al after drinking tea

SIR—It is well known that urinary aluminium excretion increases after ingestion of aluminium-containing antacid preparations¹. Discussion about the dietary intake of aluminium from tea^{2–4} would be helped by data on the bioavailability of Al known to be present in tea leaves.

We determined the aluminium content of dry, commercially available tea leaves and found it to be in the range 555–1,009 μg Al per g dry weight. In agreement with Fairweather-Tait *et al.*⁴, we find typical tea infusions (2 g tea per 150 ml tap water) contain 4.5–6.0 μg Al per ml. In contrast, typical coffee infusions contain 0.04–0.30 μg Al per ml.

To assess the bioavailability of aluminium when large quantities of tea are consumed, we monitored the urinary excretion of aluminium in six healthy male volunteers after drinking equal volumes



Urinary aluminium levels of 6 male volunteers collected after consuming tap water (left-hand column of each set), coffee (middle columns) and tea (right-hand columns) over 12 hours.

(1.2 litres) of tea, coffee or tap water on separate days. On each day participants consumed standardized meals together with either tea, coffee or tap water (300 ml) at set times. All urine passed in the 12-hour period was collected separately and assayed for aluminium by graphite furnace atomic absorption spectroscopy.

The figure shows that in every case the amount of aluminium excreted over the 12-hour period increased on the day when tea was taken. These results indicate that at least some of the aluminium present in tea is absorbed and that tea consumption must be considered in any assessment of the total dietary intake of aluminium in human beings.

KLAUS R. KOCH
M.A. BRUNO POGNET
STEPHANIE DE VILLIERS

Department of Analytical Science,
FELIX MONTEAGUDO
Department of Pharmacology,
University of Cape Town,
Rondebosch, Cape Town 7700,
South Africa

1. Kachny, W.D. *et al.* *New Engl. J. Med.* **290**, 1389 (1977).
2. Coriat, A.M. & Gillard, R.D. *Nature* **321**, 570 (1986).
3. Pearce, F. *Nat. Sci.* **114**, issue 1559, 28 (1987).
4. Fairweather-Tait, S.J. *et al.* *Nature* **330**, 213 (1987).

The nature of Taung dental maturation

SIR—In a recent letter to *Nature*¹, Conroy and Vannier describe the pattern of dental calcification and other age-related features in the immature fossil specimen from Taung (the type specimen of *Australopithecus africanus*). But their interpretation of the dental eruption and calcification sequence in the Taung child is incorrect.

The dentition of a modern human child (Fig. 1) shows a calcification sequence parallel to that of the Taung child. Most important, it possesses maxillary incisors and first molars in a state of relative development consistent with that described for the Taung specimen (Fig. 2). The minimal differences between the modern human specimen and Taung shown by Fig. 2 suggest the modern human's growth is more consistent with an ape pattern than that of the Taung specimen.

It was the relative state of development of the dentition that led Conroy and Vannier to identify Taung as ape-like, using observations on eruption patterns and on calcification stages. They observed that at eruption of the first molar (evident in both the Taung child and in the modern human shown here), no other permanent teeth are expected to have significant root development in apes. In humans, most of the anterior dentition has root forming, with about one half root calcification on the incisors (Fig. 2). Conroy and Vannier explain that these different patterns of calcification reflect the different patterns of eruption between apes and humans. According to this view, modern humans are 'unique' in that the eruption of incisors and first molars occurs together during one relatively brief time span, at about 6 years of age. This contrasts with the great-ape pattern, in which incisor eruption is "delayed for about two years after the first molar eruption"¹.

These assumptions on the relation between calcification and subsequent eruption must be questioned. The modern human (Fig. 1) has a calcification pattern

and an erupted first-molar characteristic of the proposed ape pattern, like the Taung child. This pattern is not uncommon in samples of immature human dentition from the archaeological sites of Hasanlu and Tepe Hissar that are housed at the University Museum. In the 25 specimens assessed to date, none exhibits the proposed human pattern of M1/I1

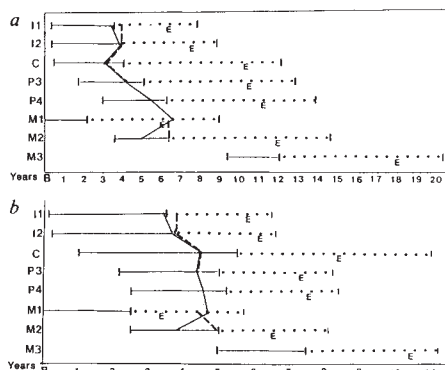


Fig. 2 Maxillary dentition of the Taung child (solid line) and a modern human child (dashed line) plotted against human² (a) and ape³ (b) standards as outlined in refs 3 and 4. The maxillary dentition is characterized by a central incisor (I1) with minimal root initiation; lateral incisors (I2) are crown complete; canines (C) are 3/4 crown complete; P3, 3/4 crown complete; P4 are absent; M1 exhibits some wear with ~ 1/2–2/3 root completion (root length ~ 7.3 mm); M2, crown completely calcified; M3 without trace. The distinctions from the Taung specimen include minimal root on central incisors (equivalent to ~ 6 months growth beyond crown complete²) and a completed M2 crown; Taung is said to be without incisor roots and M2 to be only 1/2 crown complete. The lower dentition, as Conroy and Vannier note, shows a slightly greater degree of calcification.

calcification and eruption; in fact, most show a calcification pattern of the anterior dentition in the presence of erupted first molars similar to the Taung child. Based on the eruption status of the M1 together with the calcification of the M1 root, the crown and root calcification of M2 and the

calcification status of the incisors, the incisors in these individuals would probably erupt from 2–3 years after the M1, precisely as proposed for the Taung specimen. Whether or not such a time lag would have occurred in these individuals, or in the Taung child, is unknown. It is merely suggested by a chart of the observable calcification and eruption stages (such as Fig. 2). Our sample, however, which parallels the dental development of the Taung specimen, is that of known modern humans and a conclusion that identifies them as apes is not reasonable.

The existence of known modern human children with dental calcification sequences and eruption patterns like that proposed for the Taung specimen leads to one of several conclusions: many modern humans grow in an ape-like pattern; or the patterns proposed and used here are incapable of distinguishing apes from humans and cannot be applied to characterize the nature of development in fossil hominid specimens.

ALAN MANN

Department of Anthropology,
University of Pennsylvania,
325 University Museum, Philadelphia,
Pennsylvania 19104-6398, USA

1. Conroy, G.C. & Vannier, M.W. *Nature* **329**, 625–627 (1987).
2. Moorrees, C.R.A., Fanning, E.A. & Hunt, E.E. *J. dent. Res.* **42**, 1490–1502 (1963).
3. Dean, M.C. & Wood, B.A. *Folia primatol.* **36**, 111–127 (1981).
4. Smith, B.H. *Nature* **323**, 327–330 (1986).

Orbital forcing and the Vostok ice core

SIR—In their paper on the Vostok ice core Jouzel *et al.*¹ say that the 100-kyr oscillation that dominates the climatic record over the past million years is generally assumed to result from a nonlinear response of ice sheets to orbital forcing rather than directly from eccentricity changes. We take exception to this statement and, further, we wish to point out some recent theoretical modelling results that can account for many of the observational results presented in the important follow-on Vostok CO₂ papers^{2,3}.

Although there is certainly strong evidence that orbital forcing influences the course of ice growth and decay⁴ (particularly with regard to the ~20 and 40-kyr variations), there is no general agreement that nonlinear responses of ice sheets to orbital forcing are the ultimate cause of the major 100-kyr ice age Pleistocene. On the contrary, it is also plausible to assume that we would still have the major ice ages even in the complete absence of orbital forcing. The most convincing, but by no means the only argument for this view was discussed recently in *Nature* by Mix⁵, to the effect that the 'dominant' 100-kyr oscillation of the late Pleistocene came into being as a rapid transition from an

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Fig. 1 Dentition of an immature modern human from Hasanlu, Iran (University Museum catalogue number 73-5-58), dated about 3,000 years before present.

earlier Pleistocene devoid of these major ice ages in spite of similar orbital forcing. Although not mentioned as a possibility by Mix, this state of affairs suggests that the control of these ice fluctuations may reside in the internal instability of the slow-response parts of the climatic system that allows the possibility for a bifurcation to a free, long-period, oscillatory mode. We have suggested models illustrating this possibility (see, for example, ref. 6).

The most recent version of such an 'oscillator' model⁷ emphasizes the potential role of atmospheric CO₂ as the critical free variable in the climatic system that, along with the global ice mass and some measure of the deep ocean state, can form the basis for a 'natural' nonlinear oscillator. It was shown that, by virtue of the many possibilities for positive feedback inherent in recently proposed biological, chemical and physical scenarios⁸⁻¹², the carbon cycle can provide the instability necessary to drive major climatic variability without external forcing. Moreover, with the addition of orbital forcing, a model containing relatively few adjustable parameters (rate constants) can be constructed that can give a 'first order' account of the evolution of some of the main Pleistocene variables — $\delta^{18}\text{O}$ and surface temperature⁴, global ice mass,

and the atmospheric CO₂ variations suggested by Shackleton and Pisias¹³ and now supported to a large extent by the Vostok findings¹⁻³. Although there are discrepancies, this model (which at present is the only closed, self-consistent, time-dependent model purporting to treat all these phenomena simultaneously) predicts a good deal of the variability discussed in these Vostok reports, including the phase lags.

BARRY SALTZMAN

KIRK A. MAASCH

*Department of Geology and Geophysics,
Yale University, PO Box 6666,
New Haven, Connecticut 06511, USA*

1. Jouzel, J. *et al.* *Nature* **329**, 403-408 (1987).
2. Barnola, J.M., Raynaud, D., Korotkevich, Y.S. & Lorius, C. *Nature* **329**, 408-414 (1987).
3. Genthon, C. *et al.* *Nature* **329**, 414-418 (1987).
4. Hays, J.D., Imbrie, J. & Shackleton, N.J. *Science* **194**, 1121-1132 (1976).
5. Mix, A.C. *Nature* **327**, 370 (1987).
6. Saltzman, B. & Sutera, A. *J. Atmos. Sci.* **44**, 236-241 (1987).
7. Saltzman, B. *Climate Dynam.* **1**, 77-85 (1987).
8. Broecker, W.S. *Geochim. Cosmochim. Acta* **46**, 1689-1705 (1982).
9. Keir, R.S. & Berger, W.H.J. *Geophys. Res.* **88**, 6027-6038 (1983).
10. Ennever, F.K. & McElroy, M.B. *Am. geophys. Union Geophys. Monogr.* **32**, 154-162 (1985).
11. Toggweiler, J.R. & Sarmiento, J.L. *Am. geophys. Union Geophys. Monogr.* **32**, 163-184 (1985).
12. Wenk, T. & Siegenthaler, U. *Am. geophys. Union Geophys. Monogr.* **32**, 185-194 (1985).
13. Shackleton, N.J. & Pisias, N.G. *Am. geophys. Union Geophys. Monogr.* **32**, 303-317 (1985).

The GB790325b γ -ray error box revisited

SIR—In the light of the recently discovered optical flashes¹ about 5 arc min from the published 1 arc min error box of the γ -burst (GRB) GB790325b (ref. 2), scrutiny of the γ -ray localization is called for³. We have therefore examined our previous work to see whether the flashes were produced by GB790325b. We find that the GRB error box cannot be enlarged significantly, and that an extremely improbable combination of circumstances would be required to move it a distance equivalent to several times its characteristic dimension. Thus, the optical flasher is probably either a new phenomenon or a serendipitous nearby GRB or soft γ -repeater (SGR).

GB790325b was observed by seven instruments on five spacecraft^{2,4}. The published error-box boundaries use the three 'best' data sets, with the remainder of the data providing confirmation in the form of larger, overlapping error boxes. The three crucial instruments are the Lockheed ARPA-301 experiment on the US Air Force P78-1 satellite (ARPA), the Pioneer Venus Orbiter γ -burst detector (PVO) and the Signe experiment on Venera 12 (S12). The uncertainty in the GRB position is caused by uncertainties in the relative arrival times of the γ -ray wavefront at the nominal spacecraft locations. For this event, a 100-ms timing error at any given spacecraft corresponds approximately to a 1 arc min error in the GRB location. It follows that GRB/flasher

compatibility requires arrival time uncertainties several times larger than we used previously; or unknown errors in the spacecraft ephemerides or clock calibrations; or mistakes in the calculations.

The lag uncertainty in cross-correlating data from the various experiments usually accounts for most of the timing uncertainty. Fortunately, GB790325b contained a distinct fiducial — an intense, isolated, narrow (300-ms-wide full-width-half-maximum) peak about halfway through the GRB. Thus, lag uncertainties well below 100 ms should be expected *a priori*. In fact, analyses using the entire available time histories gave 1σ uncertainties (determined using a modified χ^2 method) comparable to our time-bin widths of 12-32 ms. Also, we are confident that the lag calculations have not been biased by spectral variability. Our spectral data do not show strong evolution in this event, and our cross-correlation analysis gave no evidence for nonstatistical differences between the various time histories.

Additional arrival time uncertainties arise from imprecise knowledge of spacecraft clock calibrations and ephemerides. Based on design goals and experience with overdetermined error boxes (the 5 March event being the best example^{5,6}), these additional uncertainties are known to range from a few milliseconds for PVO and most near-Earth spacecraft, to at most a few tens of milliseconds for some

interplanetary probes. For the US spacecraft these accuracies are 'guaranteed' and should not be considered in the same light as statistical fluctuations. In any case, we are hard pressed to produce total 3σ arrival-time uncertainties greater than ~ 100 ms. (Hence the original 1 arc min error box.) Because a 5 arc min discrepancy corresponds to at least 15σ , our attention must turn from uncertainties to mistakes.

We have checked virtually every step in the process that generated this GRB error box. For example, we tracked spacecraft ephemerides and clock calibrations over long periods of time to look for irregularities. For S12 and ARPA, independent experiments on the same spacecraft gave internally consistent results. We checked the PVO ephemeris against the position of Venus using the *Astronomical Almanac*. All the actual calculations were done independently, and with identical results, at Goddard, Los Alamos and Toulouse laboratories. Finally, with one unlikely exception, even if the data from any one spacecraft were discarded, the flasher location would still be excluded by the other two. The exception involves 'throwing out' PVO (despite its accurate ephemeris and numerous clock calibrations) and invoking a source proper motion of 2 arc s yr^{-1} to the south-west. This motion is within about a factor of two of the maximum allowed by the flasher observations (D. Hartmann, in preparation).

In summary, it is extremely difficult to reconcile the optical and γ -ray error boxes. There seem to be four possibilities remaining. First, the optical flashes are spurious. This is probably excluded by the spatial coincidence of the three flashes. Second, a new kind of optical 'super-flaring' has been discovered. The possibility cannot be excluded and may have implications concerning Schaefer's flashes^{7,8}. Third, the flashes are from a serendipitous GRB. This requires a total GRB population of at least several thousand, but the inferred flasher repetition rate of about one per month seems high for a GRB. Finally, the flashes are from a serendipitous SGR. The repetition rate and temporal proximity of two of the flashes (and possibly the 20° galactic latitude?) support this conclusion, but the apparent rarity of SGRs⁹ is a problem for a serendipitous discovery in a small region of space.

J. G. LAROS

*Los Alamos National Laboratory,
Los Alamos,
New Mexico 87545, USA*

1. Hudec, R. *The Physics of Compact Objects: Theory versus Observations* COSPAR/IAU Symp. Sofia, 13-18 July 1987.
2. Laros, J.G. *et al.* *Astrophys. J.* **290**, 728-734 (1985).
3. Bignami, G.F. *Nature* **330**, 316-317 (1987).
4. Atteia, J.-L. *et al.* *Astrophys. J. Suppl.* **64**, 305-382 (1987).
5. Evans, W.D. *et al.* *Astrophys. J.* **237**, L7-L10 (1980).
6. Cline, T.L. *et al.* *Astrophys. J.* **255**, L45-L48 (1982).
7. Schaefer, B.E. *Nature* **294**, 722-724 (1981).
8. Schaefer, B.E. *et al.* *Astrophys. J.* **286**, L1-L4 (1984).
9. Atteia, J.L. *et al.* *Astrophys. J.* **320**, L105-L110 (1987).

Dressing for modern times

W.F. Bynum

The Reluctant Patron: Science and the State in Britain 1850–1920. By Peter Alter. Translated by Angela Davies. Berg, 77 Morrell Avenue, Oxford OX4 1NQ/St Martin's Press, New York: 1987. Pp.292. £30, \$50.

SCIENTISTS do not have noticeably longer memories than other people, and one of the several virtues of Peter Alter's book is its reminder that problems with governmental funding of scientific research are not new. Indeed, political pettiness in these matters may be the rule rather than the exception, at least in Britain. The title of the book is well chosen. Between about 1850 and 1920, British science put on its modern dress. In 1920, *scientia* was decently, albeit scantily, clothed; 70 years earlier, the same could not be said. During that period, a combination of private and public patronage had provided the essentials. In the end, the government's role had proved crucial: *malgré lui*, the state had become one of the main patrons of the scientific estate.

The process itself has occurred in all developed countries during the past sesquicentennial, but in very different ways. If the hallmarks of modern science are paid professionals doing the work, collaborative research and a marked tendency towards specialization, the German states set the pattern from the 1830s. At about that time, Charles Babbage and others were complaining that British science was in a state of decline, although in hindsight it is unclear what it was declining from. In 1830, it was, and had long been, 'amateur' in its orientation. This did not mean that there was no world-class British science in that age of Babbage, Faraday, Herschel and Lyell, merely that very few could expect to earn a living from it. The word 'scientist' was coined only in 1833, and even in 1860, a decade after Alter's rough starting date, only three or four of the 22 Britons elected to fellowship in the Royal Society had conventionally 'professional' scientific careers.

If Royal Society elections have become anything to go by, this conventional professional career has been an academic one, although both industry and the government itself also provided scientific jobs by 1920. Within the new nineteenth-century universities and university colleges in London, Birmingham, Manchester, Liverpool, Bristol and elsewhere, science was important and money for laboratories

and equipment came largely from private patrons, mostly from such self-made entrepreneurs as Josiah Mason, John Owens and Donald Currie. The scientific base in the universities was created with virtually no help from the state.

Welcome though it was, to many spokesmen of science private patronage seemed only a partial solution. Always the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Lord Haldane — "a true statesman of science".

German model was there, with its publicly endowed research institutes, its vigorous system of higher education, and its increasingly important science-based chemical and optical industries. As Britain began to lose out in world markets to Germany and the United States, the science lobby mounted its own campaign to try to convince the politicians that an enlightened public science policy (and a substantial capital investment) could reverse the trend. Key figures in the science lobby included a retired army officer, Colonel Alexander Strange (1818–1874) and, above all, Norman Lockyer (1836–1920), founder and first editor of *Nature*. Both *Nature* and another of Lockyer's creations, the British Science Guild (1905), provided outlets for his energetic espousal of science and its importance in national life. As Lord Kelvin had succinctly put it in the 1870s:

Investigations for which a large expenditure of money is necessary and which must be continued through long periods of years cannot be undertaken by private individuals. . . . [If] the government is well advised in respect to

science, it will be for the good of the nation that the government should make it part of its function to promote experimental investigations in science.

Despite decades of lobbying and some relatively modest gains, the scientific community began systematically to benefit from state patronage only from the turn of the century. Instances of this were the establishment of the National Physical Laboratory, which was modelled on the Physikalisch-technische Reichsanstalt in Charlottenburg and which formally opened in 1902; the amalgamation of various scientific institutions to form Imperial College London; and the rather unexpected spin-off of the 1911 National Health Insurance Act, the money to support the activities of the Medical

Research Committee (now Council). The latter was unexpected because it was an initiative of the government; in the first two instances, and more generally throughout the period, governments in Britain had at best only passively responded to repeated pleas from scientists. During the First World War state involvement in scientific research increased, with the creation of the Department of Scientific and Industrial Research, ancestor of the Science Research Council, and the advent of the research associations of the main science-based industries.

Alter's book (which is a fluent and updated translation of a monograph first published in German in 1982) throws light on these and related issues. Although based on an admirable amount of primary research, its main achievement is to synthesize the scattered and piecemeal historical and sociological literature on a subject of depressing relevance to this age of cuts and retrenchment. Nineteenth-century scientists longed for political leaders who had been trained in science. They were few and far between. Even the scientists' best friend in high places, Lord Haldane (1856–1928), was a barrister. Nevertheless, Haldane was a true statesman of science, active on its behalf and commemorated in the 'Haldane Principle' — the assumption (in Alter's words) that, in Britain, "science should determine its own research aims and methods and be free from any direct control by the executive". The new dawn has now arrived. The current prime minister has a degree in chemistry. What, one wonders, will future generations mean if they refer to a 'Thatcher Principle'? □

W.F. Bynum is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK.

Emigré's progress

Georges Charpak

De la Physique Avant Toute Chose. By Anatole Abragam. Editions Odile Jacob, 15 rue Soufflot, Paris: 1987. Pp. 378. Pbk FF180.

De la Physique Avant Toute Chose is the autobiography of a leading figure in modern physics. Anatole Abragam's contributions to nuclear magnetism have won him high distinctions, while some of his books have been standard works for decades. Then, in 1984, at an age when many physicists are ready to retire, he took pleasure in making a major contribution to the use of muons as probes for condensed matter.

The word 'pleasure' is apposite, for one of the author's designs is to impart to the reader the joy which his discoveries have brought him. He speaks of his experiments with an affection which he sometimes extends to his colleagues. A physicist reading this book will find enlightening explanations of the various milestones along the historical road of nuclear magnetism.

This, however, is only one aspect of the book. The social and scientific rise of the Russian boy, who arrived in Paris in 1925, was achieved at the cost of a difficult struggle — a struggle which did not have to be faced by those who were privileged to travel the royal road, via distinguished schools, to the few laboratories where advanced research was carried out. It was not until 1945 that, at the age of 30, Abragam finally found his place in an infant laboratory, at Saclay, where the facilities and talented young colleagues made an ambitious research programme possible. The book leads us around the twists and turns of the author's career, giving a rich account of the history of pre-war and occupied France as lived by a child who arrived there at the age of 10. Abragam was already familiar with the prose of Pushkin, Lermontov and Gogol, so it was not surprising that, after the brief period of acclimatization undergone by all immigrants, he became a highly gifted pupil at a French grammar school.

Abragam paints a flattering picture of the education that he received in Russia, and has remained devoted to Russian culture throughout his life. It was no doubt the fear of a growing policy of proletarian selectivity, which would have barred their son from pursuing advanced studies, that led his parents to emigrate. Abragam's mother was a heroic medical officer in the Red Army and his father the head of a small button factory — in other words, they were bourgeois. His mother went to France with her son and daughter while his father, a 'refusenik' ahead of his

time, did not arrive until 11 years later.

A brilliant student in quest of a career, Abragam spent 18 months "wandering around" in medical studies before taking up physics and mathematics, emerging with a degree in science. This account gives us a rather critical description of French secondary and higher education. While severely slating some teachers, the author visits abounding affection and admiration on those, few and far between, who did their job well. In 1936 there began, to quote the author, his search for a direction, a guide, and for research companions which took over ten years to find.

There were also the wartime years,

CERN

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Abragam — a taste for battle.

however, under the German occupation, when sheer survival was no mean feat. This part of the history of the occupation, even looked at from the narrow viewpoint of one person's existence, is interesting. After a year's military service, Abragam arrived at Saclay and, like many other young theoretical physicists of the time, owed the opportunity finally to meet teachers and problems equal to his ambitions to a stay abroad. His first period at the Clarendon Laboratory, Oxford, between 1948 and 1950, was to be decisive. He describes the Clarendon's fruitful environment with affection and humour, and is even now obviously still enthralled by the memory of its archaic customs. I feel that scientific historians will greatly appreciate his vivid portraits of some of the great characters in physics whom he thus met, but he also provides the ordinary reader with plethoric descriptions of scientific discoveries.

Abragam's judgements, however, are often brusque and open to question. For example, he criticizes Frédéric Joliot for failing to leave France during the war, which Joliot could easily have done, in order to continue his research in Britain or the United States. In doing so, Joliot would undoubtedly have been able to retain the leading position in nuclear

research which he had held before the war. Instead, he elected to stay in France and take part in the Resistance movement. The author may perhaps have failed properly to appreciate the fact that, for a physicist who has enjoyed all the glories of discovery and devotion to science, physics might not prevail above all else and that it might speak greatly in Joliot's favour that he could conceive a passion for a cause aimed at nothing less than re-fashioning the fate of humanity. Moreover, did not the scientific upsurge in post-war France also owe a great deal to those who, by making use of their prestige, created the structures which provided the young physicists emerging from the war with the means of obtaining essential facilities?

Abragam's sometimes cutting criticisms of certain of his contemporaries would undoubtedly have led him into a few duels at the time when academicians' blades were not worn merely as ornaments. He would nevertheless probably have taken part with enthusiasm, for his literary style often betrays a taste for battle. His sparkling exchanges with the mathematician René Thom on the respective roles of mathematics and physics are an example and a delight to read.

Readers may occasionally be confused by the wide variety of avenues along which this book leads them. They will however often smile, and sometimes laugh, at the abundance of anecdotes of which Abragam is an avid collector. So we must render thanks to the author who, via the portraits he draws of all those who have crossed his path, also gives us a glimpse of his own that goes far beyond the description that he gives of himself.

For those who may be irritated to find the events of half a century gauged by their impact on the story of an intelligent little boy from Russia, Abragam quotes in his preface this sentence uttered by a famous humorist: "False modesty is better than no modesty at all". □

Georges Charpak is at CERN, CH-1211, Geneva 23, Switzerland.

New journals review

On 29 September *Nature* will publish the eighth annual review supplement devoted to science journals. Criteria for inclusion of a journal in the 1988 issue are that:

- (i) the first number appeared, or the journal was retitled, between June 1986 and May 1987 (the second cut-off date allows at least three issues of a journal to have been published, the minimum number on which a reasonable judgement can be based);
- (ii) it is published at least three times a year;
- (iii) the main language used is English.

Publishers and learned societies are invited to send four different issues of each suitable periodical, including the first and most recent numbers (if from outside the United Kingdom, by air mail) to: The Review Editor, *Nature*, 4 Little Essex Street, London WC2R 3LF, England.

Reasons for leaving home

Paul Harvey

Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics. Edited by B. Diane Chepko-Sade and Zuleyma Tang Halpin. *University of Chicago Press*: 1987. Pp. 342. Hbk \$55, £43.95; pbk \$19.95, £15.95.

TYPICALLY, male mammals leave their place of birth in order to breed, while the female stays at home if social conditions allow. In contrast, among most bird species the females move farther than the males. The reason for these opposite patterns in the two groups is, to my mind, a mystery. Behavioural studies of known individuals from several species indicate that incest avoidance is often involved. If that is the case, only one sex need move. But why should it be males in mammals and females in birds? The various theories based on whether the breeding system is one of resource or mate defence have never convinced me, primarily because they do not explain the exceptions without a good deal of special pleading.

The facts of sex-biased and limited dispersal in mammals are both the starting point and the crux of Chepko-Sade and Halpin's edited volume. The subtitle of the book, *The Effects of Social Structure on Population Genetics*, could cover a multitude of sins. So what exactly are the editors after? If one number encapsulates their Grail it is N_e , Sewall Wright's effective or neighbourhood population size. As the editors write in their preface (p.xiv), they are interested in the usefulness of N_e for "characterizing and comparing populations and predicting the relative importance of Wright's shifting balance theory in the tempo and mode of evolution". It is not surprising that the Quest turned out to be of more value than the End.

The quality of data on dispersal emerging from long-term studies of mammals is improving. This volume brings together a representative sample of detailed descriptive but non-genetic work on known individuals from deer, horses, wolves, bears, mongooses, prairie dogs, kangaroo rats and pikas. Data from these studies are used in the final chapter to estimate N_e , which generally turns out to be surprisingly small (often tens rather than hundreds of individuals). The editors and their coauthors warn of the tentative nature of their estimates, but conclude that "the shifting balance process is likely to occur very effectively in . . . horses, . . . mongooses, . . . prairie dogs, . . . kangaroo rats and . . . pikas" (p.314). They also consider the "numerous genetic consequences" (p.314) that their analyses

imply, including the importance of genetic drift and a high rate of inbreeding.

In an earlier chapter, however, Alan Templeton challenges these results when he points out that the levels of inbreeding they imply would eliminate inbreeding depression. The near universal occurrence of inbreeding depression in studies of captive populations of mammals seems to contradict this prediction. Templeton concludes that the estimates of effective or neighbourhood population size are probably incorrect because assumptions of the models are not fulfilled. As he points out, a clearer picture will probably emerge when we have accurate estimates of paternity (genetic fingerprinting should see to that), when we find whether recolonizations following local extinctions are from one or several neighbouring groups, and when we get a clearer indication of

whether human disturbance has led to recent subdivision of populations being studied.

In addition to the chapters I have mentioned, there are others on the genetic structure of small mammal populations, on dispersal and demography in two human populations, and on the evolution of kin selection in complex groups. But it is because the editors have pushed the field data to their limits that this book is most valuable. Those with ideas about the causes and consequences of dispersal in natural populations will be helped to formulate their own questions in terms of the available data. More often, though, they will find that the necessary information has yet to be collected. □

Paul Harvey is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Fast forward

Frederick James

Collider Physics. By Vernon D. Barger and Roger J. N. Phillips. *Addison-Wesley*: 1987. Pp. 592. \$44.95, £29.95.

IT COULD be argued that collider physics as an isolated subject does not exist, and Barger and Phillips here contribute to that argument by including a considerable amount of material well outside the realm indicated by their title. The result is a relatively self-contained text on modern particle physics suitable for advanced graduate students and young research workers. Given the rapid advances in this field in recent years, especially on the theoretical side, I suspect that more than one established physicist will also benefit from the coherent (although necessarily dense) presentation and the emphasis on up-to-date topics as studied in current and near-future colliders.

The text was computer-prepared using T_eX , and is well laid out, but because of the relatively large typeface there is not as much of it as one would expect in 592 pages. The index is good enough to make the book also handy as a reference work.

One useful aspect of the book is the large number of derivations of expressions for matrix elements corresponding to particular Feynman diagrams. These are not just a help in understanding the physics, but also provide the working physicist with a good starting point for standard model calculations.

On the other hand, the experimentalist will not be fooled by the attempt to make the coverage appear comprehensive by inclusion of a few pages on detectors and the colliders themselves: the book is resolutely orientated towards the theoretical side. This is hardly surprising because the

authors are theoretical physicists, and one would not expect a complete treatment of experimental problems, but some topics of special interest to experimentalists are skipped over too lightly. For example, only a few lines are devoted to Bhabha scattering, and expressions for cross sections for polarized electrons are given without any comment on the problems of obtaining polarized beams or on the importance of polarization for the measurement of physically interesting quantities.

The chapter on Monte Carlo simulations includes listings of several Fortran programs. I applaud this growing practice — Fortran is a widely understood language and a good vehicle for the precise expression of certain ideas, becoming in fact the scientific equivalent of a French or Latin quotation in a history book. Readers should not however take literally the authors' claim that the programs are written according to the Fortran 77 standard, because they contain some constructs (such as lower-case characters and `REAL*8` declarations) which are not part of the standard and are not accepted by several popular compilers.

Barger and Phillips are, of course, at their best in the chapters devoted to the more speculative theoretical developments: entire chapters are devoted to heavy quark production, the Higgs boson(s), the fourth generation and the various higher symmetries. This is after all what collider physics is all about, and we have here about as clear a snapshot of this fast-moving subject as the circumstances allow. In these chapters the authors have chosen to concentrate on the most promising classes of theories rather than particular models. We should know within a few years whether their choice was judicious. □

Frederick James is in the Data Handling Division at CERN, 1211 Geneva 23, Switzerland.

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information
about the Clearinghouse,
please fill out and mail back
the coupon below.

UMI Article
Clearinghouse

Yes! I would like to know more about UMI
Article Clearinghouse. I am interested in
electronic ordering through the following
system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL
Subsystem

- ☐ Other (please specify) _____
☐ I am interested in sending my order by
mail.
☐ Please send me your current catalog and
user instructions for the system(s) I
checked above.

Name _____
Title _____
Institution/Company _____
Department _____
Address _____
City _____ State _____ Zip _____
Phone (_____) _____

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Literary life

Richard Mabey

The Book of Naturalists: An Anthology of the Best Natural History. Edited by William Beebe. Princeton University Press: 1988. Pp. 499. Hbk \$45, £25.20; pbk \$12.50, £7.

TOWARDS the end of this "Anthology of the Best Natural History", there is a piece by Aldous Huxley's friend Gerald Heard, in which he tries to give a kind of report from the front on the discovery of flint tools. His jumpy, inquisitive "half-man" finds the nodules sticking out of a cave wall, scares himself silly striking sparks and eventually succeeds in splitting one open:

In the light he saw no spark. But it had been angry. When he looked at the nodule's end a corner had been bitten off. It was black inside. It had a faint smell. He licked the smooth black piece, then dropped it angrily. It had bitten him. His tongue was cut.

Enter proto-Stukeley, man the romantic geologist, alongside man the tool-maker.

Rather sadly there isn't much of this kind of blend of fascination and animism, which was later to produce the most imaginative writing on the natural world. William Beebe (1877-1962), who compiled the first edition of this classic collection in 1944, took a strictly rational and developmental view of natural history. Those first stone-agers may have set (or broken) the mould; but it was their intelligence, not their feelings, that was important. Since then, according to Beebe, there has been an ever continuous progression in increase of knowledge, an evolution worthy of comparison with that other evolution of animal life on our planet — the foundation of all our labors and our love of science.

Beebe charts this intellectual progress from Aristotle through Darwin to Rachel Carson, and lists the "ideal equipment" for a writer of literary natural history: enthusiasm, tempered with infinite patience and an absolute devotion to truth; keen senses; opportunity for observation; "thorough training in laboratory technique"; and so on.

With these criteria there is no room for writers such as Richard Jefferies and John Clare, exceptional observers who were nonetheless not prepared to see our relationships with nature in purely scientific terms. But it is a rich and intriguing anthology for all that, especially when its contributors are shown with their soft parts exposed. There is Frederick II explaining bird migration in 1245 — a remarkable insight not least because these vast, intricately navigated flights must have seemed so much less reasonable than, say, hibernation; Leeuwenhoek becoming quite soppy about the "little animals"

bustling about under his microscope; Linnaeus off-duty, describing the sloth in a footnote: it "turns its head as though in astonishment; call, an exciting sensarius; noise frightful, tears pitiful".

Charles Waterton's 1825 essay on the same animal is ponderous and affected by comparison; and the editor's preference, when it comes to the point, for the literal rather than the literary, means that there are also plenty of examples of the more pompous and domineering face of scientific natural history: Audubon omniscient on wild turkeys; Theodore Roosevelt preaching the masculine virtues of wilderness; Edward Armstrong patronizing the black guillemot ("a quaint, jolly, little fellow"). John Muir's lyrical account of the water ouzel (1894), in which he imagines the bird as an extension of its white-water habitat, is much better. But pushing the image as far as he can (the bird's form is "as smoothly compact as a pebble that has been whirled in a pot-hole") he obliterates the real bird by the metaphor.

Problems of this kind are endemic to nature writing. There are just no easy ways across the species barrier as there are in other kinds of literature. Empathy too easily degenerates into anthropomorphism. Minutely detailed objective descriptions can take the subject out of the land of the living altogether. The one constant is the observers themselves, and it may be that the proper — and certainly the most honest — subject of literary nature writing is not a biological 'object' but an encounter.

There is one outstanding example of this in Beebe's anthology, a moving account (1932) by Gustav Eckstein of some days shared with two pet rats. It is an extraordinary piece, affectionate, intuitive, piercingly observed, yet refusing to be contained within purely human terms of reference — nature writing for a world in which the kindredness of life has been realized.

Towards ten every evening the two take turns to bathe. I have fitted a board across the basin under the tap where the water drips one drop at a time. To be wet all over makes them very weak and very unhappy, but to catch one drop, and wash vigorously with that, and then catch another, that is different. I myself look forward to it — to see the way they rise from the board, put out those marvelous hands, and wait for the drop. □

Richard Mabey, 10 Cedar Road, Berkhamsted, Hertfordshire HP4 2LA, UK, is author, among other books, of Gilbert White: Biography of the Author of *The Natural History of Selborne* (Century, 1986).

● On the one hundredth anniversary of the birth of Aldo Leopold, the American conservationist, Oxford University Press have issued a commemorative edition of his *A Sand County Almanac*, first published in 1949. The book describes the changes in the Wisconsin countryside over a year. Price is \$17.95, £12.95.

Roles of G protein subunits in transmembrane signalling

Eva J. Neer* & David E. Clapham*†

* Department of Medicine, Cardiovascular Division, Brigham and Women's Hospital Harvard Medical School, Boston, Massachusetts 02115, USA

† Departments of Pharmacology, Physiology and Biophysics, Mayo Foundation, Rochester, Minnesota 55905, USA

A family of proteins called G proteins couples cell surface receptors to a variety of enzymes and ion channels. Since many cells contain several very similar G proteins, an important question is how signals remain specific as they cross the cell membrane.

CELLS communicate with each other by a variety of signal molecules, such as neurotransmitters, hormones and growth factors, that are detected by specific receptors on the plasma membrane of the responding cell. Stimulation of a receptor by its agonist initiates a cascade of biochemical processes that produce an intracellular signal, which ultimately causes a change in the behaviour of the cell—for example, secretion of an enzyme, contraction or initiation of cell division.

One class of receptors, including the phototransducing molecule rhodopsin, the muscarinic acetylcholine receptor and the β -adrenergic receptor, operates by activating a membrane-bound transducing protein termed a GTP-binding or G protein (Fig. 1). There are several known G proteins, regulating different intracellular pathways that result in changes in the activity of specific effectors, such as ion channels in the plasma membrane, adenylyl cyclase or enzymes of phosphatidyl inositol metabolism.

A current problem in understanding the operation of these transmembrane signalling systems is to explain how specificity is achieved. Although there is surprising sequence conservation among the G protein linked receptors¹, the activation of each receptor by its agonist is highly specific. On the other hand, the next steps in signal transmission, activation of G proteins and subsequent modulation of membrane effectors, seem less precise, with several receptors often regulating each intracellular pathway through one or more very similar G proteins²⁻¹³. For example, in the heart, the acetylcholine and adenosine receptors both activate the same K^+ channel¹⁴. Possible interactions among receptors, G proteins and effectors are shown in Fig. 2. The question we will examine here is how, given these very similar and sometimes interchangeable proteins, the cell might specify the range of responses to hormones. We will focus particularly on the role of the G proteins linked to hormone and neurotransmitter receptors. We propose that the structural similarity among the components makes it likely that the stoichiometry and the physical arrangement of the components in the cell membrane will be important for the ability of the cell to segregate and respond correctly to external stimuli.

Function and organization

The G proteins are composed of three subunits, α , β and γ , and are associated with the plasma membrane. Figure 1 summarizes the current view of the mechanism of activation of G proteins. Excellent recent reviews give a detailed description of this process^{15,16}.

Figure 3 illustrates the possible arrangement of the signalling proteins in the plasma membrane, conforming as much as possible to what is known about the physical arrangement of hormone receptors, G proteins and effectors in the lipid bilayer. The G proteins are shown entirely on the cytoplasmic face of the membrane. The exact arrangement of the G protein subunits is not known so we have shown three possibilities. Although

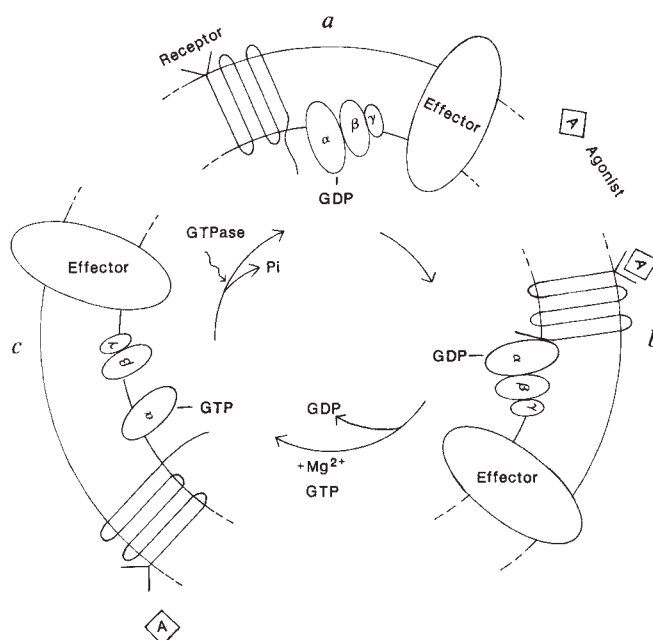


Fig. 1 Simple scheme for G protein signal transduction. From the unliganded state (a), receptor binds agonist A (for example, epinephrine; acetylcholine) which produces a change (b) in receptor-G protein interaction, allowing GTP, in the presence of Mg^{2+} , to replace GDP on the α subunit. The activated α -GTP subunit and the $\beta\gamma$ subunits dissociate and one or both interacts with effectors (for example, adenylyl cyclase, K^+ channel). Alternatively, free $\beta\gamma$ may bind other α subunits. The intrinsic GTPase activity of the α subunit hydrolyses GTP to GDP, releasing inorganic phosphate (P_i), and α -GDP recombines with $\beta\gamma$ (c), ending the activation cycle^{15,16}. Nonhydrolysable analogues of GTP such as Gpp(NH)p or GTP- γ S produce persistent activation of α subunits and persistent dissociation of α from $\beta\gamma$ because activation cannot be reversed by hydrolysis of these nucleotide analogues to GDP.

detergents are required to solubilize the hormone receptor-linked G proteins from the plasma membrane, once solubilized, the α subunits behave as hydrophilic molecules^{17,18}. This suggests that $\beta\gamma$ subunits may anchor the α subunits to the membrane¹⁹, although α subunits probably also interact with the membrane directly. If $\beta\gamma$ subunits were the only attachment, activation of the heterotrimer by GTP- γ S (a non-hydrolysable analogue of GTP which causes persistent dissociation of subunits in solution^{15,16}) would be expected to release α subunits into the cytoplasm. In general, this has not been observed, although it may occur in some cells²⁰. Buss *et al.* have recently shown that α subunits which are substrates for the enzymic activity of pertussis toxin are myristoylated, but those that are

cholera toxin substrates are not, nor is the $\beta\gamma$ complex²¹. Other membrane proteins known to be myristoylated⁴ may bind to the cytoskeleton²².

The $\beta\gamma$ component must form a single functional unit, as native β and γ do not dissociate (E.J.N., unpublished). The $\beta\gamma$ subunits vary in hydrophobicity from very hydrophobic (brain $\beta\gamma$; ref. 17) to hydrophilic (red cell $\beta\gamma$; ref. 18). This variable hydrophobicity is not due to the amino acid sequence of the β subunit, as the sequences deduced from complementary DNA for two forms of β subunit neither predict a hydrophobic protein nor suggest any membrane-spanning regions²³⁻²⁷. Therefore, either covalent modifications of β or variations in the γ subunit create the observed differences in hydrophobicity. The scheme in Fig. 3 shows three possible arrangements of the $\alpha\beta\gamma$ heterotrimer which agree with available structural information.

One consequence of the arrangements suggested in Fig. 3 is that all the important actions of the α and $\beta\gamma$ subunits take place in the cytoplasmic compartment, even though the α subunits do not dissociate from the plasma membrane. *In vitro*, the resolved adenylyl cyclase catalytic unit combines with G protein in aqueous solution^{28,104}, and G proteins can modify receptor function without a bilayer²⁹. But all successful reconstitutions of the complete hormone-responsive system have been in lipid vesicles (see references to Table 2). The lipid bilayer may only be needed to orient the components since the probability of all components (receptor, G protein and effector) aligning correctly in solution may be low.

Heterogeneity of α subunits

Bacterial toxins from *Vibrio cholera* or *Bordetella pertussis* can covalently modify the G proteins by the addition of an ADP-ribose group to the α subunit. Both toxins are enzymes which transfer ADP-ribose from NAD to specific acceptor sites on the G proteins^{30,31,96,101}. Susceptibility to ADP-ribosylation divides the known G proteins into four broad groups: substrates for cholera toxin only, pertussis toxin only, both toxins or neither toxin^{15,16}. The latter group is as yet uncharacterized. Initially, the cholera toxin substrates were thought only to stimulate and the pertussis toxin substrates only to inhibit adenylyl cyclase. They were therefore named G_s and G_i respectively. It is now apparent that both of these categories of α subunits have a broader functional repertoire, including regulation of ion transport and phosphatidyl inositol metabolism (reviewed in refs 15, 16). Nevertheless it is still convenient to call the cholera toxin substrates α_s -like and the pertussis toxin substrates α_i -like, as structural analysis justifies their division into two families. The third category of G proteins, the transducins of retinal cones and rod outer segments, couples light activated rhodopsin to a phosphodiesterase and is a substrate for both toxins.

The cloning efforts of the past two years have revealed a large number of new α and β subunits of G protein with surprising sequence similarities within the set of α and β proteins. The similarity of their sequences has made cloning of closely related cDNAs easier but has made separation of the proteins and correlation of proteins with gene products harder. Because sequence data for G proteins have become available before function has been defined, we suggest that the subunits should be identified by their structure, rather than their assumed function, as the latter practice is leading to considerable confusion.

The G proteins are a sampler of the means nature can use to generate protein diversity. There are at least four variants of α_s with only modest functional differences^{33,34} which are related by alternative splicing of messenger RNA, probably from a single gene^{12,13}. Molecular cloning has identified the four types of actual or proposed pertussis toxin substrates (apart from the retinal transducins) which are called α_0 , α_{i-1} , α_{i-2} and α_{i-3} (refs 6-11). These forms do not result from alternative mRNA splicing, but are products of separate genes located on different chromosomes³⁵. Proteins corresponding to these genes have been definitively identified only for α_{i-1} and α_0 (the 41K and

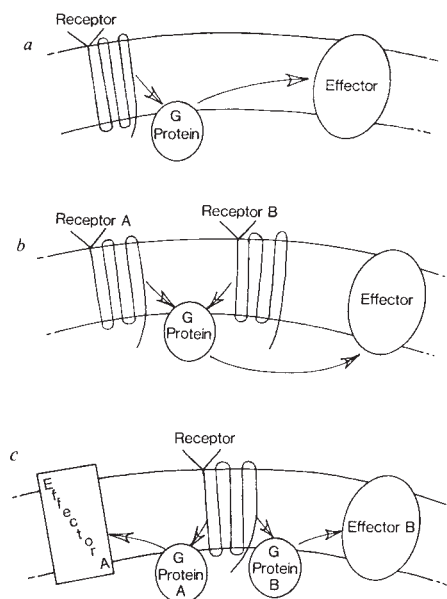


Fig. 2 Possible interactions among components of the signalling system. *a*, Completely specific: one kind of agonist bound receptor activates one specific G-protein which activates one kind of effector. By a single kind of receptor we mean one gene product; the four subtypes of muscarinic receptor are considered to be four independent receptors. By effectors we mean enzymes or ion channels which are directly regulated by the G protein, for example, adenylyl cyclase, phospholipases A_2 and C, K^+ channels, Ca^{2+} channels. As with receptors, there are likely to be several isoforms of each effector but at present little is known about effector subtypes. *b*, Specific for effector: a single G protein can be activated by more than one receptor and activates a single effector. *c*, Multiple interactions: a single receptor activates two or more G proteins, each of which activate an effector. Other permutations are possible.

39K pertussis toxin substrates from brain^{6,7,8,36,37}); α_0 is structurally similar to α_{i1-3} and can, therefore, be considered another α_i -like protein.

A very important distinction between the transducins of retinal rods and cones and other G proteins is that the transducin α subtypes are segregated into different cells³⁸, unlike the hormone receptor-linked α subunits where individual cell types contain at least four or five different α subunits¹¹. In addition to α_s , which is universally present, some but not all single cell types contain mRNA for all three α_i -like proteins¹¹. Many kinds of bone-marrow-derived cells lack α_{i-1} ¹¹. However, other α subunits may be found some of which, like transducin, will be cell-specific.

Structure

Because many laboratories independently set out to clone G protein from different species, it quickly became possible to compare the structures of the subtypes across species. The α subunit subtypes are highly conserved, with interspecies differences in amino acid sequence of <3%. Such conservation suggests that each subtype has a specific role in signal transduction. Figure 4 shows a comparison of α_0 with one of the α_i subtypes (α_{i-3}). The most conserved regions, indicated by the solid blocks at the top of the figure, are the regions thought to make up the guanine nucleotide binding site on the basis of their similarity to the oncogene *ras* proteins and bacterial elongation factor Tu (EF-Tu)^{39,40,93-95}. The amino acid differences which distinguish one subtype from another are not located at random in the sequence of each α subunit but are clustered in well-defined regions^{11,39}. A particularly prominent cluster of variability occurs between amino acids 85 and 125. Although this region is different from one α subunit to another, it is highly

conserved in the same α subtype across species. Masters *et al.*³⁹ initially proposed that this region might be involved in effector interactions, and tested their hypothesis by constructing a chimaeric α protein containing the amino terminal 60% of α_i and the carboxyl terminal 40% of α_s . The chimaeric molecule activated adenylyl cyclase, indicating that the region between residues 85 and 125, which came from an inhibitory α subunit, was not sufficient to determine effector specificity⁴¹.

If this region (residues 85–125) does not specify which effector is regulated by the α subunit, what other function might it have? It may be important in targeting α subunits to membrane domains, either by interacting directly with cytoskeletal elements or perhaps with other proteins ('molecular chaperones')⁴² whose function is to ensure that proteins assemble correctly into macromolecular complexes. Alternatively, the region may interact with as yet undefined proteins, modulating G protein function. Other functional regions of the α subunits are indicated at the bottom of Fig. 4.

Heterogeneity of $\beta\gamma$ subunits

The $\beta\gamma$ subunits are also heterogeneous, but this heterogeneity may arise in part through stable combinations of similar β subunits with different γ subunits. Cells other than retinal cells contain two closely-related β subunits, a major 36K form (β_{36}) and a minor 35K form (β_{35})^{43,44}. Evans *et al.*⁴⁴ have defined two types of nonretinal γ subunits. As β can be dissociated from γ only by denaturing agents, the existence of two β and at least two γ subunits allows at least four stable β - γ combinations in nonretinal cells. The combinations may be functionally different. Retinal cells contain only β_{36} , whose deduced amino acid sequence is identical to β_{36} from nonretinal cells^{23–27}. The γ subunit associated with retinal β_{36} is immunologically different from that associated with nonretinal β_{36} (refs 43, 105).

Reconstitution of G proteins

Because individual cells contain several closely-related α and $\beta\gamma$ subunits, we need to understand the minimal structural requirements that determine which G proteins interact with which receptors or effectors. Reconstitution of purified G proteins with pure receptors or with membranes is one way to determine which associations are possible. The results show considerable flexibility in the interaction of receptors and effectors with G proteins.

Purified heterotrimers. Table 1 lists functions reconstituted with pure G protein heterotrimers. Coupling of a receptor with a G protein affects receptor affinity for agonist, the GTPase activity of the G protein, and GDP/GTP exchange. The interaction of

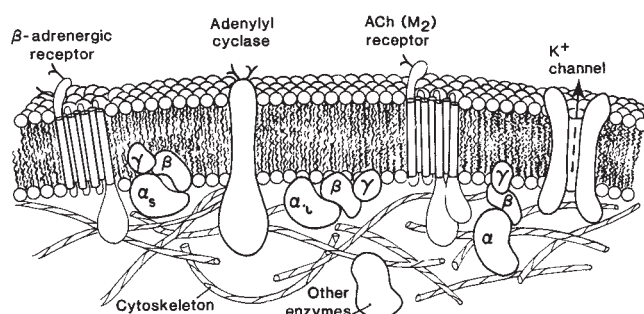


Fig. 3 Schematic view of the organization of receptors, G proteins and effectors in the plasma membrane. All receptors so far isolated are transmembrane glycoproteins. Receptors that work through the G proteins such as rhodopsin, β -adrenergic receptors, and muscarinic receptors have a similar structure which appears to include seven transmembrane helices and regions of homology in the cytoplasmic loops¹. Adenylyl cyclase is a transmembrane glycosylated enzyme whose active site is on the cytoplasmic side of the membrane⁹¹. The structure of the muscarinic K^+ channel is not yet known although other ion channels are complex multi-meric proteins⁹². The structure of the G proteins is discussed in the text. Three possible arrangements of α and $\beta\gamma$ subunits are illustrated. The α and/or $\beta\gamma$ subunits may interact with the cytoskeleton^{21,22,86–90}.

G protein with effectors changes enzymatic or ion channel activity. Effectors should reciprocally modify G protein function but have not yet been shown to do so.

In phospholipid vesicles as well as in cells, the same G protein can interact with several receptors. More than fifteen distinct kinds of receptors can stimulate adenylyl cyclase through G_s . Conversely, a single receptor can interact with more than one G protein. The interpretation of chemical reconstitution data is made somewhat harder by the recent demonstration of multiple receptor subtypes. A different kind of reconstitution is achieved by expression of cDNA for a single muscarinic cholinergic receptor subtype in a cell lacking muscarinic receptors. Results of such experiments suggest that even a single receptor subtype can interact with more than one G protein⁴⁵.

In some cases, G proteins seem to interact more specifically with effectors than with receptors. For instance, only G_s (and not G_i or G_o) can stimulate adenylyl cyclase. Table 1 suggests that in the G_i -like family, interchange of G proteins at the effector is possible as both G_o and G_i reconstitute the increase in phosphoinositol metabolism stimulated by *f*-Met-Leu-Phe. The range of heterogeneity within the G_i -like proteins was not known when many of the experiments in Table 1 were done, however. A more precise evaluation of the homogeneity of G_i and G_o proteins will probably reveal that the 'pure' proteins are mixtures of subtypes and that some G protein-effector crosstalk is only apparent. But even if there were an exact correspondence between G protein and effector, if the receptor-G protein interface is not specific, each receptor could set loose a shower of activated G proteins. The end result would be indiscriminate activation of response pathways. Therefore, specificity of G protein-effector interaction is not, in itself, sufficient to prevent mixed signals.

Isolated α and $\beta\gamma$ subunits. So far, we have considered the G proteins as a single unit. But the hallmark of all known receptor-linked G proteins is that when the $\alpha\beta\gamma$ heterotrimer binds a nonhydrolysable analogue of GTP in solution, it dissociates into α and $\beta\gamma$ subunits^{15,16}. Assuming that receptor-stimulated GTP binding also causes dissociation of the subunits in the membrane (which has not yet been demonstrated) then both subunits become potential signal molecules. Table 2 summarizes the known actions of isolated α or $\beta\gamma$ subunits.

Table 1 Functions of pure, activated heterotrimers

Heterotrimer	Target	Reference
$\alpha_s\beta\gamma$	β -Adrenergic receptor	46, 47
	Adenylyl cyclase catalytic unit	18, 28, 48–50, 69
$\alpha_i\beta\gamma$	β -Adrenergic receptor ($\alpha_i\beta\gamma < \alpha_s\beta\gamma$)	51, 55
	Rhodopsin ($\alpha_i\beta\gamma = \alpha_T\beta\gamma$)	54, 55
	GABA receptor	53
	α_2 -Adrenergic receptor	52
	Muscarinic cholinergic receptor	56, 57
	<i>f</i> -Met-Leu-Phe-stimulated phosphoinositol production	58
$\alpha_o\beta\gamma$	Adenylyl cyclase catalytic unit	59, 60
	GABA receptor	53
	α_2 -Adrenergic receptor	52
	Muscarinic cholinergic receptor	56, 57
	<i>f</i> -Met-Leu-Phe-stimulated phosphoinostol production	58
$\alpha_T\beta\gamma_T$	Rhodopsin	54, 55
	β -Adrenergic receptor ($\alpha_T\beta\gamma_T < \alpha_s\beta\gamma$)	55
	G_s -stimulated adenylyl cyclase catalytic unit	60

Table 2 Functions of isolated α or $\beta\gamma$ subunits

Subunit	Target	Reference
α_s	Adenylyl cyclase catalytic unit	50
	Calcium channel	61
α_i^*	Rhodopsin	54
	Adenylyl cyclase catalytic unit	62, 70
	cAMP phosphodiesterase	63
	K ⁺ channel	64, 73
α_0	Rhodopsin	67
	Muscarinic cholinergic receptor	56
	K ⁺ channel	73
$\beta\gamma$	K ⁺ channel	65
	Phospholipase A ₂	66
	cAMP phosphodiesterase	63
	G _s -stimulated adenylyl cyclase catalytic unit	33, 69, 71
	Unstimulated adenylyl cyclase catalytic unit	68

We have included only reconstitutions with pure components or membrane fragments and not reconstitutions by injection of subunits into whole cells¹⁰³.

* The relative content of α_{i1-3} is not known. This designation includes α_i from human erythrocytes now called α_k by Codina *et al.*⁶⁴.

Results of many experiments demonstrating α subunit function have been generalized to indicate that the α subunits are the active members of all heterotrimers. Pure, activated α_s , free of $\beta\gamma$, stimulates the isolated catalytic unit of adenylyl cyclase⁵⁰ and increases the molecular weight of the catalytic unit by about 40K, consistent with binding of an α_s subunit⁷². A 41K α_i protein can inhibit the adenylyl cyclase catalytic unit, although two other α_i -like α subunits cannot^{50,62,70}. The transducin α subunit (α_T) activates retinal cGMP phosphodiesterase¹⁵. Codina *et al.*⁶⁴ reported activation of a cardiac K⁺ channel by a 40K α_i -like protein isolated from erythrocytes. This observation has been reproduced in our laboratory. In addition, we find that α_0 from bovine brain is as effective as the erythrocyte protein in activating the muscarinic gated K⁺ channel⁷³.

Isolated α subunits can also interact with hormone receptors. For example, Florio and Sternweis⁵⁶ reconstituted α_0 subunits and muscarinic cholinergic receptors into phospholipid vesicles and showed that receptors can increase the GTPase activity of the α_0 subunit. Some hormone receptors, however, may interact with both subunits since the transducin $\beta\gamma$ ($\beta\gamma_T$) subunit is required to couple the α subunit of transducin (α_T) to rhodopsin¹⁵ and $\beta\gamma$ enhances the effect of α_0 on the muscarinic receptor⁵⁶. A binding site for $\beta\gamma$ on a receptor has not yet been directly demonstrated.

One role of the $\beta\gamma$ subunit is to attenuate the action of the α subunit⁵⁰. Smigel *et al.*⁵⁰ reasoned that since the GTP-ligand α subunit is an active species, increasing the concentration of $\beta\gamma$ should favour reassociation and hence deactivation. They and other laboratories tested and confirmed the hypothesis^{50,69,71}.

Although $\beta\gamma$ reverses the action of α , more recent experiments demonstrate that $\beta\gamma$ itself can modulate the function of effectors (see Table 2). The $\beta\gamma$ subunits activate retinal phospholipase A₂ (ref. 66), and inhibit brain cAMP phosphodiesterase⁶³ besides the indirect inhibition of adenylyl cyclase by interaction with α_s , described above, $\beta\gamma$ inhibits unstimulated brain adenylyl cyclase either directly or via interaction with calmodulin⁶⁸.

In cardiac atrial myocytes, $\beta\gamma$ activates the muscarinic-gated K⁺ channel ($i_{K_{ACh}}$), an effect which is reversed by preincubation of $\beta\gamma$ with unactivated α_{39} and α_{41} subunits⁶⁵. The $\beta\gamma$ subunits purified from bovine brain and from human placenta, but not from retina, are equally effective⁷³. Careful controls have shown that activation is specific to $\beta\gamma$. Logothetis *et al.*⁷³ report activation of the channel at $\beta\gamma$ concentrations of 0.2–1 nM. Y. Kurachi, I. Katada and M. Ui (personal communication) have activated $i_{K_{ACh}}$ in guinea pig atria with 10 pM $\beta\gamma$.

The K⁺ channel of cardiac atrial myocytes seems to be an example of an effector regulated both by activated α and by $\beta\gamma$

subunits^{64,65,73}. Both α and $\beta\gamma$ subunits activate the same K⁺ channel ($i_{K_{ACh}}$, conductance 40 pS, open time 1 ms) in chick embryonic, rat neonatal and guinea pig atria, although the GTP- γ S-liganded α subunit generally does so at lower concentrations than $\beta\gamma$ (between 1 pM α_{40} (ref. 64) and 10 pM α_{40} or α_{39} (ref. 73)). The significance of differences in the concentration dependence of subunit action is hard to evaluate, as the α protein is hydrophilic while $\beta\gamma$ is hydrophobic. The α subunit may be able to reach the channel directly from the aqueous phase, whereas $\beta\gamma$ may need to insert into the membrane. Both subunits may be important *in vivo*. For example, two separate binding sites in the channel might explain the biexponential desensitization kinetics of the K⁺ channel. Defining the individual functions of the two subunits will require further study not only of the G proteins but of the structure of the K⁺ channel itself.

Models of G protein action

The overall structural similarity within the α and the $\beta\gamma$ subunit families, the *in vitro* ability of some subunits to exchange for others, and the multiplicity of cellular functions carried out by each class of subunits raise questions about the mechanisms insulating hormonal signals from excessive crosstalk. The most common view, which we term the free dissociation model, is that hormone receptors, G protein heterotrimers, their dissociated subunits and effector molecules can all float freely in the membrane. Any subunit can bind any receptor or effector if it has the right structural specificity.

There is good evidence that some components of the response system can move in the plane of the lipid bilayer. Hormone-sensitive adenylyl cyclase can be reconstituted by fusing cells which contain either receptors or adenylyl cyclase catalytic units⁷⁴ indicating that components from the two types of cells must have diffused in the plane of the plasma membrane. A similar conclusion follows from kinetic analysis of inactivation of β -adrenergic receptors in turkey erythrocytes which suggests that receptors can couple to many adenylyl cyclase catalytic units⁷⁵. Two independent receptors (adenosine and acetylcholine) have been shown to share the same pool of cardiac K⁺ channels in a single cell¹⁴ and convergence of receptors on a single pool of adenylyl cyclase molecules^{76,77} has also been demonstrated. Unless the receptors are adjacent to each other, all these observations suggest that the components do move in the membrane.

The issue of whether receptors, G proteins or effectors can move in the membrane is separate from the question of whether dissociated G protein subunits diffuse away from each other after activation by agonist-liganded receptors. No direct evidence demonstrates that dissociation of the heterotrimer occurs in the membrane, but two indirect experiments do indicate that dissociation can occur. For example, excess $\beta\gamma$ can inhibit adenylyl cyclase in membranes^{78,79}, suggesting that added $\beta\gamma$ associates with free α_s in the membrane and shifts the subunit equilibrium. Pure unliganded α_i , the putative inhibitory α subunit, actually activates adenylyl cyclase in platelet and S49 cell membranes, presumably by binding free $\beta\gamma$ subunits^{78,79,102}.

If all components are entirely free to float and if there is potential crosstalk at the receptor-G protein or G protein-effector interfaces, then the stoichiometry of receptors, G proteins and effectors becomes an extremely important determinant of specificity. For example, although β -adrenergic agonists release $\beta\gamma$ from G_s, they do not activate K⁺ channels in the heart probably because the concentration of G_s is usually substantially lower than G_i and too little $\beta\gamma$ would be released from G_s to be significant. Moreover, all α subunits do not have the same affinity for $\beta\gamma$; some may dissociate more readily than others⁸⁰. Therefore, one pool of $\alpha\beta\gamma$ heterotrimers may be a less effective source of $\beta\gamma$ than another. Finally, receptor number can affect the amount of G protein activated.

Another way to ensure specificity of G protein action is to

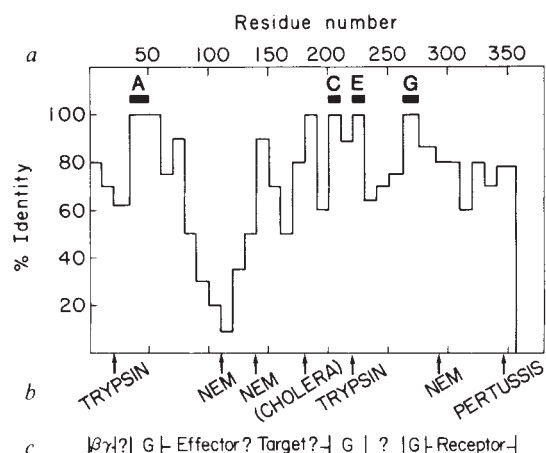


Fig. 4 Structural and functional features of α subunits. Section *a* illustrates the point that two pertussis toxin substrates (α_{i-3} and α_0) are identical in regions thought to make up the guanine nucleotide binding site (A, C, E, G) and quite different around amino acid 100 and at the carboxyl terminus. In the figure, the number of identical amino acid residues in a given segment is expressed as a percentage and plotted as a function of the position of the segment. The length of the segment (11 ± 3 residues) was chosen to maximize the identity. The boxes marked A, C, E, G indicate regions where the α subunits are similar to the GTP-binding region in *ras* and EFTu^{39,40,93,94}. A similar plot comparing α_{i-1} or α_{i-2} to α_{i-3} shows that identities and differences occur in similar places in all these α subunits¹¹. The observation that regions of maximal difference do not occur randomly along the polypeptide chain suggests that the differences may be functionally important. *b*, Summary of chemically reactive sites on the α_0 (α_{39} subunit). Analysis of the functional consequences of the modifications has provided clues to the potential roles of regions of α_0 in signal transduction. Sites of tryptic cleavage were identified by Hurley *et al.*⁹⁵. Removal of the amino-terminal 2K tryptic fragment interferes with the ability of α_0 and α_1 to associate with $\beta\gamma$ ¹⁰⁰. Cholera toxin is in parentheses because it has not yet been shown to modify the α_{39} protein. But the arginine that is ADP-ribosylated by cholera toxin in transducin¹⁰¹ is present in α_{39} and the sequences surrounding this residue are conserved among α subunits. Pertussis toxin has been shown to ADP-ribosylate residue 347 in transducin⁹⁶. The α_{39} protein contains the same cysteine and is highly homologous to transducin in this area⁶. ADP-ribosylation does not prevent dissociation of α_{39} from $\beta\gamma$ ⁸⁰. The cysteines modified by *N*-ethylmaleimide are at the arrows marked NEM. The reactivity of cysteine 108 is not affected by GTP- γ S; the other two sulphhydryls are blocked by GTP γ S. Modification of Cys 108 blocks ADP-ribosylation of Cys 347 by pertussis toxin but does not prevent association of α_0 with $\beta\gamma$ ⁹⁷. *c*, Putative functional regions of α_0 . The assignments are partly based on the chemical studies and partly on an analogy to *ras* or EFTu (see ref. 39). Since the protein is folded, the real structure of functional regions will be complex and unlikely to be divided so neatly into domains. For example, reaction of Cys 108 in the 'effector' domain profoundly affects the function of residues in the 'receptor' domain⁹⁷.

constrain the location of the receptor, G protein, and/or effector. We call this the constrained dissociation model. Within a limited area, subunits could dissociate, but they would not be free to diffuse throughout the membrane. Several lines of evidence suggest that such organization may exist.

Dufau *et al.*⁸¹ showed that cAMP generated by adenylyl cyclase molecules coupled to the gonadotropin receptor is directed to a small pool of protein kinase molecules, suggesting compartmentation of the proteins involved in hormone responses. The cAMP made in response to isoproterenol in cultured heart cells can also be distinguished from that generated in response to prostaglandin E_1 (ref. 82). Neutron-inactivation analysis of hormone-sensitive adenylyl cyclase indicates that receptors and G proteins form multi-enzyme arrays, and it has

been suggested that the signalling complex is an aggregate larger than the hormone receptor, G protein and adenylyl cyclase put together^{83,84}. But other studies have found smaller target sizes in other cell types⁸⁵. Finally, there is evidence that the cytoskeleton influences the ability of hormones to stimulate adenylyl cyclase and of detergents to solubilize G proteins⁸⁶⁻⁸⁹, and that the $\beta\gamma$ subunit appears to associate with cytoskeletal components⁹⁰. Furthermore, the observation that myristylated *src* protein may be associated with the cytoskeleton²² suggests a function for the recently described myristoylation of α_0 and α_1 ²¹. If hormone receptors, G proteins and effectors indeed exist in specialized membrane domains then it will be important to define the structures that target the molecules to their common meeting ground and keep them there.

The two models are not mutually exclusive and each may even apply to different parts of the cell. Showing that particular receptors, G proteins and effectors segregate to a particular surface of a cell, the apical rather than the basolateral for example, might appear to argue against the free dissociation model. In some epithelial cells such localization has long been known, as parathyroid hormone-sensitive and basal adenylyl cyclase are found preferentially in the basolateral surface of renal cortical cells⁹⁸. More recently Dominguez *et al.*⁹⁹ have shown that although adenylyl cyclase is localized to the basolateral surface in intestinal epithelia, G_s seems to predominate in the apical surface. Nevertheless, there may be free diffusion of components within the basolateral surface. In more symmetrical cells, the regions of constrained diffusion might be confined to small patches of membrane.

Conclusion

The nature of a cell's response to hormones will depend on its endowment of receptors, G proteins and effectors, as well as on the relative concentrations and affinities of all the components. In addition, some geographical organization may help to maintain the specificity of cellular responses to hormones. To understand signal transduction, we must define the general principles governing the function of transducing molecules. Some of the broad outlines are now understood. The next challenge is to explain the unique functions of differentiated cells. Here, a variety of mechanisms may be used to turn the general mechanisms to particular tasks. These considerations make a starting point for future experiments to determine the precise stoichiometry of the components in cells, to define the structural basis for constraints on the motion of these membrane-associated molecules and to discover whether additional cell components are involved in modulating the transmission of hormonal signals across the plasma membrane of cells.

Many of the ideas and analyses in the present review were foreseen by M. Rodbell in his insightful article in *Nature*¹⁰⁶ written before the complexity of receptor and G protein families had been defined. Some were raised in a pithy discussion by H. R. Bourne¹⁰⁷.

This work was supported by the NIH. We thank Ms Barbara Zillman and Mrs Paula McColgan for expertly typing the manuscript.

1. Dohlman, H. G., Caron, M. G. & Lefkowitz, R. J. *Biochemistry* **26**, 2657-2664 (1987).
2. Lochrie, M. A., Hurley, J. B. & Simon, M. *Science* **228**, 96-99 (1985).
3. Tanabe, T. *et al. Nature* **315**, 242-245 (1985).
4. Medynski, D. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 4311-4315 (1985).
5. Yatsunami, K. & Khorana, H. G. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4316-4320 (1985).
6. Itoh, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3776-3780 (1986).
7. Michel, T., Winslow, J. W., Smith, J., Seidman, J. G. & Neer, E. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7663-7667 (1986).
8. Nukada, T. *et al. FEBS Lett.* **197**, 305-310 (1986).
9. Sullivan, K. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 6687-6691 (1986).
10. Jones, D. T. & Reed, R. R. *J. biol. Chem.* **262**, 14241-14249 (1987).
11. Kim, S. Y. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
12. Robishaw, J. D., Smigel, M. D. & Gilman, A. G. *J. biol. Chem.* **261**, 9587-9590 (1986).
13. Bray, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 8893-8897 (1986).
14. Kurachi, Y., Nakajima, T. & Sugimoto, T. *Pflügers Arch. ges. Physiol.* **407**, 264-274 (1986).
15. Stryer, L. & Bourne, H. R. *A. Rev. Cell. Biol.* **2**, 391-419 (1986).
16. Gilman, A. G. *A. Rev. biochem.* **56**, 615-650 (1987).
17. Huff, R. M., Axton, J. M. & Neer, E. J. *J. biol. Chem.* **260**, 10864-10871 (1985).

18. Codina, J. *et al. J. biol. Chem.* **259**, 5871-5886 (1984).
19. Sternweis, P. C. *J. biol. Chem.* **261**, 631-637 (1986).
20. Rodbell, M. *Trends Biochem. Sci.* **10**, 461-464 (1985).
21. Buss, J. E., Mumby, S. M., Casey, P. J., Gilman, A. G. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
22. Hamaguchi, M. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2312-2316 (1987).
23. Sugimoto, K. *et al. FEBS Lett.* **191**, 235-240 (1985).
24. Fong, H. K. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 2162-2166 (1986).
25. Codina, J., Stengel, D., Woo, S. L. & Birnbaumer, L. *FEBS Lett.* **207**, 187-192 (1986).
26. Fong, H. K. W., Amatruda, T. T., Birren, B. W. & Simon, M. I. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3792-3796 (1987).
27. Gao, B., Gilman, A. G. & Robishaw, J. D. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6122-6125 (1987).
28. Strittmatter, S. & Neer, E. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6344-6348 (1980).
29. Limbird, L. E., Gill, D. M. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 775-779 (1980).
30. Gill, D. M. & Meren, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3050-3054 (1978).
31. Katada, T. & Ui, M. *J. biol. Chem.* **257**, 7210-7216 (1982).
32. Sternweis, P. C., Northup, J. K., Smigel, M. D. & Gilman, A. G. *J. biol. Chem.* **256**, 11517-11526 (1981).
33. Northup, J. K., Smigel, M. D., Sternweis, P. C. & Gilman, A. G. *J. biol. Chem.* **258**, 11369-11376 (1983).
34. Graziano, M. P., Casey, P. J. & Gilman, A. G. *J. biol. Chem.* **262**, 11375-11381 (1987).
35. Neer, E. J., Michel, T., Eddy, R., Shows, T. & Seidman, J. G. *Hum. Genet.* **77**, 259-262 (1987).
36. Neer, E. J., Lok, J. M. & Wolf, L. G. *J. biol. Chem.* **259**, 14222-14229 (1984).
37. Sternweis, P. C. & Robishaw, J. J. *J. biol. Chem.* **259**, 13806-13813 (1984).
38. Lerea, C. L., Somers, D. E., Hurley, J. B., Klock, I. B. & Bunt-Milam, A. M. *Science*, **234**, 77-80 (1986).
39. Masters, S. B., Stroud, R. M. & Bourne, H. R. *Protein Engng* **1**, 47-54 (1986).
40. Halliday, K. R., *J. cyclic Nucleotide Protein Phosphor Res.* **9**, 435-448 (1984).
41. Masters, S. B., Sullivan, K. A., Miller R. T., Beiderman, B., Lopez, N. G., Ramachandran, J. & Bourne H. R. (in preparation).
42. Ellis, J. *Nature* **238**, 378-379 (1987).
43. Roof, D. J., Applebury, M. L. & Sternweis, P. C. *J. biol. Chem.* **260**, 16242-16249 (1985).
44. Evans, T., Fawzi, A., Fraser, E. D., Brown, M. L., Northup, J. K. *J. biol. Chem.* **262**, 176-181 (1987).
45. Ashkenazi, A. *et al. Science* **238**, 672-675 (1987).
46. Pedersen, S. E. & Ross, E. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7228-7232 (1982).
47. Asano, T., Pedersen, S. E., Scott, C. W. & Ross, E. M. *Biochemistry* **23**, 5460-5467 (1984).
48. May, D. C., Ross, E. M., Gilman, A. G. & Smigel, M. D. *J. biol. Chem.* **260**, 15829-15833 (1985).
49. Cerione, R. A. *et al. J. biol. Chem.* **259**, 9979-9982 (1984).
50. Smigel, M. *et al. Adv. cyclic Nucl. Prot. Phosphoryl Res.* **17**, 1-18 (1984).
51. Asano, T., Katada, T., Gilman, A. G. & Ross, E. M. *J. biol. Chem.* **259**, 9351-9354 (1984).
52. Kim, M. H. & Neubig, R. R. *Biochemistry*, **26**, 3664-3672 (1987).
53. Asano, T., Ui, M. & Ogasawara, N. *J. biol. Chem.* **260**, 12653-12658 (1985).
54. Kanaho, Y. *et al. J. biol. Chem.* **259**, 7378-7381 (1984).
55. Cerione, R. A. *et al. J. biol. Chem.* **260**, 1493-1500 (1985).
56. Florio, V. A. & Sternweis, P. C. *J. biol. Chem.* **260**, 3477-3483 (1985).
57. Kurose, H. *et al. J. biol. Chem.* **261**, 6423-6428 (1986).
58. Kikuchi, A. *et al. J. biol. Chem.* **261**, 11558-11562 (1986).
59. Hildebrandt, J. D., Codina, J. & Birnbaumer, L. *J. biol. Chem.* **259**, 13178-13185 (1984).
60. Cerione, R. A. *et al. Biochemistry* **24**, 4499-4503 (1985).
61. Yatani, A. *et al. Science* **238**, 1288-1291 (1987).
62. Katada, T., Oinuma, M. & Ui, M. *J. biol. Chem.* **261**, 5215-5221 (1986).
63. Asano, T., Ogasawara, N., Kitajima, S. & Sano, M. *FEBS Lett.* **203**, 135-138 (1986).
64. Codina, J., Yatani, A., Grenet, D., Brown, A. M. & Birnbaumer, L. *Science* **236**, 442-445 (1987).
65. Logothetis, D. E., Kurachi, Y., Galper, J., Neer, E. J. & Clapham, D. E. *Nature* **325**, 321-326 (1987).
66. Jelsema, C. L. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3623-3627 (1987).
67. Tsai, S.-C., Adamik, R., Kanako, Y., Halpern, J. L. & Moss, J. *Biochemistry* **26**, 4728-4733 (1987).
68. Katada, T., Kusakabe, K., Oinuma, M. & Ui, M. *J. biol. Chem.* **262**, 11897-11900 (1987).
69. Neer, E. J., Wolf, L. G. & Gill, D. M. *Biochem. J.* **241**, 325-336 (1987).
70. Katada, T., Oinuma, M., Kusakabe K. & Ui, M. *FEBS Lett.* **213**, 353-358 (1987).
71. Cerione, R. A. *et al. Biochemistry* **26**, 1485-1491 (1987).
72. Bender, J. L., Wolf, L. G. & Neer, E. J. *Adv. cycl. Nucl. Prot. Phosph. Res.* **17**, 101-110 (1984).
73. Logothetis, D. E., Kim, D. H., Northup, J., Neer, E. J. & Clapham, D. E. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
74. Schramm, M., Orly, J., Eimerl, S. & Korner, M. *Nature* **268**, 310-313 (1977).
75. Tolkovsky, A. M. & Levitzki, A. *Biochemistry* **17**, 3811-3816 (1978).
76. Birnbaumer, L. & Rodbell, M. *J. biol. Chem.* **244**, 3477-3482 (1969).
77. Sabol, S. L. & Nirenberg, M. *J. biol. Chem.* **254**, 1913-1920 (1979).
78. Katada, T., Bokoch, G. M., Northup, J. K., Ui, M. & Gilman, A. G. *J. biol. Chem.* **259**, 3568-3577 (1984).
79. Katada, T., Northup, J. K., Bokoch, G. M., Ui, M. & Gilman, A. G. *J. biol. Chem.* **259**, 3578-3585 (1984).
80. Huff, R. M. & Neer, E. J. *J. biol. Chem.* **261**, 1105-1110 (1986).
81. Dufau, M. L. *et al. J. biol. Chem.* **253**, 3721-3729 (1978).
82. Buxton, K. O. & Branton, L. L. *J. biol. Chem.* **258**, 10233-10239 (1983).
83. Schlegel, W., Kempner, E. S. & Rodbell, M. *J. biol. Chem.* **254**, 5168-5176 (1979).
84. Martin, B. R., Stein, J. M., Kennedy, E. L., Doberska, C. A. & Metcalfe, J. C. *Biochem. J.* **184**, 253-260 (1979).
85. Skorecki, K., Verkman, A. S., Jung, C. Y. & Ausiello, D. A. *Am. J. Physiol.* **250**, C115-C123 (1986).
86. Sahyoun, N. E., LeVine, H., III, David, J., Hebbon, G. M. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6158-6162 (1981).
87. Cherskey, B. C., Zadunaisky, J. A. & Murphy, R. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6401-6405 (1980).
88. Insel, P. A. & Kennedy, M. S. *Nature* **273**, 471-473 (1978).
89. Rasenick, M. M., Stein, P. J. & Bitensky, M. L. *Nature* **294**, 560-562 (1981).
90. Carlson, K. E., Woolkalis, M. J., Newhouse, M. G. & Manning, D. R. *Molec. Pharmacol.* **30**, 463-468 (1986).
91. Smigel, M. *J. biol. Chem.* **261**, 1976-1982 (1986).
92. Stevens, C. *Nature* **328**, 198-199, 1987.
93. Jurnak, F. *Science* **230**, 32-36 (1985).
94. la Cour, T. F. M., Nyborg, J., Thirup, S. & Clark, B. F. C. *EMBO J.* **4**, 2385-2388 (1985).
95. Hurley, J. B., Simon, M. I. & Teplow, D. B. *Science* **226**, 860-862 (1984).
96. West, R. E., Moss, J., Vaughan, M., Liu, T. & Liu, T. Y. *J. biol. Chem.* **260**, 14429-14430 (1985).
97. Winslow, J. W., Bradley, J. D., Smith, J. A. & Neer, E. J. *J. biol. Chem.* **262**, 4501-4507 (1987).
98. Shlatz, L. J., Schwartz, I. L., Kinne-Saffran, E. & Kinne, R. *J. mem. Biol.* **24**, 131-144 (1975).
99. Dominguez, P., Velasco, G., Barros, F. & Lazo P. S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6965-6969 (1987).
100. Neer, E. J., Pulsifer, L. & Wolf, L. G. *J. biol. Chem.* (in the press).
101. Van Dop, C., Tsubokawa, M., Bourne, H. R. & Ramachandran, J. *J. biol. Chem.* **259**, 696-698 (1984).
102. Katada, T., Bokoch, G. M., Smigel, M. D., Ui, M. & Gilman, A. G. *J. biol. Chem.* **259**, 3586-3595 (1984).
103. Hescheler, J., Rosenthal, W., Trautwein, W. & Schultz, G. *Nature* **325**, 445-447 (1987).
104. Ross, E. M. *J. biol. Chem.* **256**, 1949-1953 (1981).
105. Gierschik, P., Codina, J., Simons, C., Birnbaumer, L. & Spiegel, A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 727-731 (1985).
106. Rodbell, M. *Nature* **284**, 17-22 (1980).
107. Bourne, H. R. *Nature* **325**, 296-297 (1987).

ARTICLES

Quantitative assessment of worldwide contamination of air, water and soils by trace metals

Jerome O. Nriagu* & Jozef M. Pacyna†

*National Water Research Institute, Box 5050, Burlington, Ontario L7R 4A6, Canada

†Norwegian Institute for Air Pollution Research, PO Box 64, N-2001 Lillestrom, Norway

Calculated loading rates of trace metals into the three environmental compartments demonstrate that human activities now have major impacts on the global and regional cycles of most of the trace elements. There is significant contamination of freshwater resources and an accelerating accumulation of toxic metals in the human food chain.

THE inventory of emissions from different industrial sources is needed both in global mass balance models for trace metals and also for relating the mesoscale variations in aerosol concentration and composition to global circulation patterns. Estimates of the source strengths are a necessary first step in the design of pollution control programmes, and are also invaluable in the

assessment of the long-term ecological and health impacts of the large quantities of toxic metals now being dispersed globally in the different environmental compartments¹. One of us (J.O.N.) previously presented a global inventory of anthropogenic emissions of Cd, Cu, Ni, Pb and Zn to the atmosphere in 1975 (ref. 2). The present report provides a

Table 1 Emission factors for the release of trace metals to the atmosphere

Source category	Unit	As	Cd	Cr	Cu	Hg	In	Mn	Mo
Coal combustion									
-electric utilities	$\mu\text{g MJ}^{-1}$	15-100	5-25	80-500	60-200	10-35		70-450	15-150
-industry and domestic	g t^{-1}	0.2-2.10	0.1-0.5	1.7-12	1.4-5	0.5-3.0		1.5-12	0.4-2.5
Oil combustion									
-electric utilities	$\mu\text{g MJ}^{-1}$	1.0-5.0	4-30	15-100	60-400			10-100	10-70
-industry and domestic	g t^{-1}	0.02-0.2	0.05-0.2	1.0-5.0	0.5-3.0			1.0-5.0	0.3-1.5
Pyrometallurgical non-ferrous metal production									
-mining	g t^{-1} metal	5.0-10.0	0.1-0.5		20-100			50-100	
-Pb production	g t^{-1} metal	200-400	10-50		60-80	2-4.0			
-Cu-Ni production	produced	1,000-1,500	200-400		1,700-3,600		1.0-4.0	100-500	
-Zn-Cd production		50-150	200-1,000		50-150	8-45	0.5-1.0		
Secondary non-ferrous metal production	g t^{-1} waste		2.5-4.0		50-150				
Steel and iron manufacturing	g t^{-1} steel	0.5-3.5	0.04-0.4	4.0-40.0	0.2-4.0			1.5-40.0	
Refuse incineration									
-municipal	g t^{-1} waste	1.1-2.8	0.4-10	0.7-7.0	7.0-14	1.0-15		1.8-9.0	
-sewage sludge		5.0-10	1.0-12	50-150	10-60	5-20		50-100	
Phosphate fertilizers	g t^{-1} fertilizer		0.5-2.0		1.0-5.0				
Cement production	g t^{-1} cement	0.2-1.0	0.01-0.60	1.0-2.0					
Wood combustion	g t^{-1} wood	0.1-0.5	0.1-0.3		1.0-2.0	0.1-0.5			

Source category	Unit	Ni	Pb	Sb	Se	Sn	Tl	V	Zn
Coal combustion									
-electric utilities	$\mu\text{g MJ}^{-1}$	90-600	50-300	10-50	7-50	10-50	10-40	20-300	70-500
-industry and domestic	g t^{-1}	2.0-15.0	1.0-10.0	0.2-1.5	0.8-2.0	0.1-1.0	0.5-1.0	1.0-10.0	1.5-12.0
Oil combustion									
-electric utilities	$\mu\text{g MJ}^{-1}$	60-2,500	40-300		6-50	60-400		1,200-9,000	30-220
-industry and domestic	g t^{-1}	20-80	2.0-6.0		0.3-1.5	0.8-10.0		60-200	1.0-7.0
Pyrometallurgical non-ferrous metal production									
-mining	g t^{-1} metal	~100	500-1,000	1.0-10.0	1.0-2.5				50-100
-Pb production	g t^{-1} metal	85	3,000-8,000	50-100	10-50				50-120
-Cu-Ni production	produced	900	1,300-2,600	50-200	50-150	50-200		5-10.0	500-1,000
-Zn-Cd production			1,200-2,500	10-20	20-50				100,000-180,000
Secondary non-ferrous metal production	g t^{-1} waste		50-800	1-5	1-5				300-1,600
Steel and iron manufacturing	g t^{-1} steel	0.05-10.0	1.5-20.0	0.005-0.1	0.001-0.003			0.1-2.0	10-45
Refuse incineration									
-municipal	g t^{-1} waste	0.7-3.0	10-20	3.0-6.0	0.2-0.5	1.0-10			20-60
-sewage sludge		10-50	80-100	5-20	1.0-10	5.0-20		3.0-20	50-150
Phosphate fertilizers	g t^{-1} fertilizer	1.0-5.0	0.4-2.0		0.003-0.009				10-50
Cement production	g t^{-1} cement	0.1-1.0	0.02-16.0				3.0-6.0		2.0-20.0
Wood combustion	g t^{-1} wood	1.0-3.0	2.0-5.0						2.0-10.0

A blank space in this table (and Table 2) denotes an insignificant contribution from a particular source.

revision of the earlier data and extends the calculations to many more trace elements. We also present, for the first time, worldwide inventories of industrial/municipal discharges of trace metals into soils and the aquatic ecosystems. The calculations provide some perspective on the problem of toxic metal pollution as a global and regional issue.

Source function

The emission factors for the release of trace elements to the atmosphere are shown in Table 1, and are based on the review of emission studies in Western Europe, the United States, Canada and the Soviet Union³. It is known¹ that pollution control strategies in the developing countries are often less stringent than those of Europe and North America. The emission factors used in this report may thus under-represent the global rate of metal emissions.

In most cases, the ranges in the emission factors listed in Table 1 fall within a factor of 2-10. Basically, the range is determined by (1) the concentrations of the trace elements in the raw material; (2) the production technology employed in the emitting industry; and (3) the type and efficiency of the pollution control installations. The concentrations of trace elements in industrial raw material and the associated airborne wastes can obviously vary by more than a factor of 2-10. For example, the As concentrations in coal range from 0.34 to 130 $\mu\text{g g}^{-1}$ and reach 1,500 $\mu\text{g g}^{-1}$ in some Czechoslovakian

lignites⁴. Such coals with extreme As levels are used locally for domestic purposes and although they can be excluded in deriving the global As emission inventory, they certainly should be considered in estimating the local or even national emissions.

Special attention has been given to deriving the correct emission factors for various production technologies within the same industry. This is particularly true of the high-temperature processes employed in non-ferrous metal smelters (roasting, smelting and refining steps), iron and steel production (electric arc and basic oxygen furnaces versus the older open-hearth plants), and wet versus dry kiln operations in the cement industry. Refuse incineration is becoming a very important source of trace metals in the atmosphere⁴. Because of the large difference in the chemical make-up of the refuse inputs in various countries, it is difficult to select a reasonable range of emission factors for this source and the values used in this study are very tentative. Metal applications in various industries as well as specific uses of certain metals can also emit significant amounts of trace metals; such contributions have been lumped under the miscellaneous heading (Table 2).

Practically every industry discharges one trace metal or the other into the water or soil. We have limited our inventory to the principal industrial and commercial users of water and producers of solid wastes. Extensive data bases currently exist on the trace metal concentrations in industrial and municipal solid wastes and aqueous effluents (for example, see refs 5-19)

Table 2 Worldwide emissions of trace elements to the atmosphere in 1983 (10^3 kg yr^{-1})

Source category	Global production/consumption (10^9 kg yr^{-1})	As	Cd	Cr	Cu	Hg	In	Mn	Mo
Coal combustion									
-electric utilities	[$15.5 \times 10^9 \text{ MJ}$]	232-1,550	77-387	1,240-7,750	930-3,100	155-542		1,080-6,980	232-2,320
-industry and domestic	990	198-1,980	99-495	1,680-11,880	1,390-4,950	495-2,970		1,485-11,880	396-2,480
Oil combustion									
-electric utilities	[$5.8 \times 10^9 \text{ MJ}$]	5.8-29	23-174	87-580	348-2,320			58-580	58-406
-industry and domestic	358	7.2-72	18-72	358-1,790	179-1,070			358-1,790	107-537
Pyrometallurgical non-ferrous metal production									
-mining*		40.0-80	0.6-3		160-800			415-830	
-Pb production	3.9	780-1,560	39-195		234-312	7.8-16			
-Cu-Ni production	8.5	8,500-12,750	1,700-3,400		14,450-30,600	37-207	8.5-34.0	850-4,250	
-Zn-Cd production	4.6	230-690	920-4,600		230-690		2.3-4.6		
Secondary non-ferrous metal production†			2.3-3.6		55-165			1,065-28,400	
Steel and iron mfg.	710‡	355-2,480	28-284	2,840-28,400	142-2,840				
Refuse incineration									
-municipal	140§	154-392	56-1,400	98-980	980-1,960	140-2,100		252-1,260	
-sewage sludge	3	15-60	3-36	150-450	30-180	15-60		5,000-10,000	
Phosphate fertilizers	137		68-274		137-685				
Cement production	890	178-890	8.9-534	890-1,780					
Wood combustion	600¶	60-300	60-180		600-1,200	60-300			
Mobile sources	647 (gasoline)								
Miscellaneous		1,250-2,800							
Total, emissions		12,000-25,630	3,100-12,040	7,340-53,610	19,860-50,870	910-6,200	11-39	10,560-65,970	793-5,740
Median value		18,820	7,570	30,480	35,370	3,560	25	38,270	3,270
Source category	Ni	Pb	Sb	Se	Sn	Tl	V	Zn	
Coal combustion									
-electric utilities	1,395-9,300	775-4,650	155-775	108-775	155-755	155-620	310-4,650	1,085-7,750	
-industry and domestic	1,980-14,850	990-9,900	198-1,480	792-1,980	99-990	495-990	990-9,900	1,485-11,880	
Oil combustion									
-electric utilities	3,840-14,500	232-1,740		35-290	348-2,320		6,960-52,200	174-1,280	
-industry and domestic	7,160-28,640	716-2,150		107-537	286-3,580		21,480-71,600	358-2,506	
Pyrometallurgical non-ferrous metal production									
-mining*	800	1,700-3,400	18-176	18-176				310-620	
-Pb production	331	11,700-31,200	195-390	195-390				195-468	
-Cu-Ni production	7,650	11,050-22,100	425-1,700	427-1,280	425-1,700		43-85	4,250-8,500	
-Zn-Cd production		5,520-11,500	46-92	92-230				46,000-82,800	
Secondary non-ferrous metal production†		90-1,440	3.8-19	3.8-19				270-1,440	
Steel and iron mfg.	36-7,100	1,065-14,200	3.6-7.1	0.8-2.2			71-1,420	7,100-31,950	
Refuse incineration									
-municipal	98-420	1,400-2,800	420-840	28-70	140-1,400			2,800-8,400	
-sewage sludge	30-180	240-300	15-60	3-30	15-60		300-2,000	150-450	
Phosphate fertilizers	137-685	55-274		0.4-1.2				1,370-6,850	
Cement production	89-890	18-14,240				2,670-5,340		1,780-17,800	
Wood combustion	600-1,800	1,200-3,000						1,200-6,000	
Mobile sources		248,030¶							
Miscellaneous		3,900-5,100						1,724-4,783	
Total, emissions	24,150-87,150	288,700-376,000	1,480-5,540	1,810-5,780	1,470-10,810	3,320-6,950	30,150-141,860	70,250-193,500	
Median value	55,650	332,350	3,510	3,790**	6,140	5,140	86,000	131,880	

* The following primary production figures from the ores were used in the calculations: $8.0 \times 10^9 \text{ kg yr}^{-1}$ for Cu, $3.4 \times 10^9 \text{ kg yr}^{-1}$ for Pb, $6.2 \times 10^9 \text{ kg yr}^{-1}$ for Zn, and $8.3 \times 10^9 \text{ kg yr}^{-1}$ for Mn.

† The following secondary production figures were also used: $1.1 \times 10^9 \text{ kg yr}^{-1}$ for Cu, $1.8 \times 10^9 \text{ kg yr}^{-1}$ for Pb and $0.9 \times 10^9 \text{ kg yr}^{-1}$ for Zn.

‡ The value represents the production of crude steel and is used because all emission factors are calculated with reference to the production of one tonne of crude steel.

§ This figure represents 25% of the municipal refuse generated annually (see Table 4 below).

|| We have assumed that only 10% of the sewage sludge produced is incinerated.

¶ See ref. 13.

¶ It has been calculated assuming that 45% of total leaded motor gasoline in the world has 0.15 g a Pb content of 1 l^{-1} and the rest contains 0.40 g l^{-1} . Also it was assumed that Q gasoline = 0.75 kg dm^{-3} .

** This figure is for particulate Se only. Because volatile Se accounts for about 40% of the Se released²⁶, the total Se emission is estimated to be 6,320 tonnes yr^{-1} .

and the emission factors in Table 3 are based on a critical survey of the published literature. In general, where the reported concentrations appear to be too high, the lower end of the concentration range has been adopted in this report.

The average serviceable lives of the major metal-containing products are unknown. In estimating the loadings into the soils, we have assumed, quite tentatively, that for the metals (namely Mn, Mo, Ni, Sb and V) used primarily in manufacturing durable goods, only 1-5% of their global production is wasted (discarded, applied or washed off due to corrosion) annually on land. For Cd, Cu, Pb Cr and Zn which find significant applications as fertilizer, pigment, lubricant or chemicals, the wastage rate is assumed to be 5-10% of the annual production figure. For Hg (used extensively as pesticide) and Se (widely used as an additive in animal and poultry feeds), the wastage rate is conservatively estimated to be 10-15%. About 80-90% of the As produced each year is applied on soils as agricultural organic pesticides.

It is impossible to place an error range on the calculated inventories. The exact global value of the metal consumed or waste generated are unlikely to differ from those used in this report by more than a factor of two or so; a single number rather than a range in values for the annual global discharge or production/consumption has been employed in this paper. The principal uncertainty in calculating the contribution from each source therefore stems from the wide range in metal concentrations in the discharges. We have endeavoured to use the common ranges in the reported concentrations and unlike most of the previously reported inventories²⁰⁻²⁶, we have eschewed the use of an 'average' emission factor in the calculations.

Validation of the global and regional inventories of trace metal emissions published by us has been encouraging. The trace metal profiles (or records) in the Arctic snowfields²⁷, lake sediments and peats (see ref. 28 for a good overview), and soils^{5,29} are in reasonable accord with the calculated historical changes in rates of anthropogenic emissions to the atmosphere².

Table 3 Emission factors for the release of trace metals to the soil and water

Source category	As	Cd	Cr	Cu	Hg	Mn	Mo	Ni	Pb	Sb	Se	V	Zn
Water (ng l⁻¹)													
Domestic wastewater													
-Central	0.02-0.09	0.002-0.02	0.09-0.4	0.05-0.2	0-0.002	0.2-0.9	0-0.03	0.1-0.6	0.01-0.08	0-0.03	0-0.05	0-0.3	0.1-0.5
-Non-central	0.02-0.12	0.005-0.02	0.1-0.7	0.07-0.5	0-0.007	0.5-1.5	0-0.03	0.2-0.8	0.01-0.08	0-0.03	0-0.05	0-0.3	0.1-0.6
Steam electric	0.4-0.12	0.001-0.04	0.5-1.4	0.6-3.8	0-0.6	0.8-3.0	0.01-0.2	0.5-3.0	0.04-0.2	0-0.06	1.0-5.0	0-0.1	1.0-5.0
Base metal mining and dressing	0.04-1.5	0.001-0.6	0.008-1.4	0.2-18	0-0.3	1.5-23	0.005-1.1	0.02-1.0	0.5-5.0	0.08-0.7	0.5-5.0	0-0.02	0.04-12
Smelting and refining													
-Iron & steel						2.0-5.2			0.2-0.4				0.8-3.5
-Non-ferrous metals	0.5-6.4	0.004-1.8	1.5-10	1.2-8.5	0.001-0.002	1.0-7.5	0.003-0.2	1.0-12	0.5-3.0	0.04-3.8	1.5-10	0-0.6	1.0-10
Manufacturing processes													
-Metals	0.01-0.06	0.02-0.07	0.6-2.3	0.4-1.5	0-0.03	0.1-0.8	0.02-0.2	0.008-0.3	0.1-0.9	0.1-0.6	0-0.2	0-0.03	1.0-5.5
-Chemicals	0.12-1.4	0.02-0.5	0.5-4.8	0.2-3.6	0.004-0.3	0.4-3.0	0-0.6	0.2-1.2	0.08-0.6	0.004-0.5	0-0.07	0.04-1.0	
-Pulp and paper	0-0.3	—	0.004-0.5	0.01-0.13	—	0.01-0.15	—	0.02-0.04	0.004-0.3	0.001-0.09	0.002-0.3	—	0.03-0.5
-Petroleum products	0.002-0.2	0-0.04	0.001-0.7	0.002-0.2	0-0.08	—	—	0.004-0.2	0.003-0.4	—	0.002-0.3	—	0.01-0.8
Soils (μg g⁻¹)													
Agric. & food wastes	0-0.4	0-0.2	0.3-6.0	0.2-2.5	0-0.1	1.0-7.5	0.6-2.0	0.4-3.0	0.1-1.8	0-0.6	0-0.5	0.2-1.5	0.8-10
Animal wastes, manure	0.6-2.2	0.1-0.6	5.0-30	7.0-40	0-0.1	25-70	2-12	1.5-18	1.6-10	0-0.43	0.2-0.7	1.0-5.5	75-160
Logging and other wood wastes	0-0.3	0-0.6	0.2-1.6	0.3-4.7	0-0.2	1.6-9.5	0-0.3	0.2-2.1	0.6-7.5	0-0.5	0-0.3	0.1-0.9	1.2-15
Urban refuse	0.2-1.6	2.0-17	15-75	30-90	0-0.6	55-320	0.5-10	5.0-23	40-150	0.3-5.0	0.1-1.5	0.2-1.2	80-220
Municipal sewage sludge	0.3-12	1.0-20	8-550	240-1,030	0.5-9.0	220-540	4.0-16	25-110	140-480	2.1-10	0.3-6.9	11-73	900-2,800
Miscellaneous organic wastes including excreta	0-0.25	0-0.06	0.04-2.3	0.2-2.9	0-0.02	0.4-3.0	0.3-1.9	0.8-15	0.08-7.6	0-0.5	0-0.4	0.5-3.6	0.6-10
Solid wastes, metal mfg.	0.03-0.6	0-0.2	1.7-6.3	2.5-20	0-0.1	1.1-13	0.03-0.4	2.2-6.5	11-28	0-0.4	0.03-0.5	0.1-0.6	7.0-50
Coal fly ash and bottom ash†	1.8-10	0.4-3.6	40-120	25-90	0.1-1.3	134-445	4.1-20	15-75	12-65	0.7-6.0	1.1-16	3.0-18	30-130
Fertilizer	0-0.1	0.2-15	0.2-2.3	0.3-3.5	0-0.02	0.8-5.0	0-0.1	1.2-3.3	2.5-14	0-0.03	0.1-0.6	0.2-0.8	1.6-6.5
Peat (agricultural and fuel uses)	0.1-1.3	0-0.3	0.1-0.5	0.4-5.2	0-0.05	14-45	0.4-2.0	0.6-9.4	1.2-6.8	0.1-1.2	0-1.1	0.2-4.5	0.4-9.4

* The emission factors shown, which correspond to the common concentrations of trace metals in the solid wastes or liquid effluents have been compiled from a wide variety of sources including refs 5-19.

† The concentrations given are for the coal rather than for fly ash or bottom ash.

Table 4 Anthropogenic inputs of trace metals into the aquatic ecosystems (10⁶ kg yr⁻¹)

Source category	Annual global discharge (10 ⁹ m ³)	As	Cd	Cr	Cu	Hg	Mn	Mo	Ni	Pb	Sb	Se	V	Zn
Domestic wastewater†														
-Central	90	1.8-8.1	0.18-1.8	8.1-36	4.5-18	0-0.18	18-81	0-2.7	9.0-54	0.9-7.2	0-2.7	0-4.5	0-2.7	9.0-45
-Non-central	60	1.2-7.2	0.3-1.2	6.0-42	4.2-30	0-0.42	30-90	0-1.8	12-48	0.6-4.8	0-1.8	0-3.0	0-1.8	6.0-36
Steam electric	6	2.4-14	0.01-0.24	3.0-8.4	3.6-23	0-3.6	4.8-18	0.1-1.2	3.0-18	0.24-1.2	0-0.36	6.0-30	0-0.6	6.0-30
Base metal mining and dressing	0.5	0-0.75	0-0.3	0-0.7	0.1-9	0-0.15	0.8-12	0-0.6	0.01-0.5	0.25-2.5	0.04-0.35	0.25-1.0	—	0.02-6
Smelting and refining														
-Iron and steel	7						14-36			1.4-2.8				5.6-24
-Non-ferrous metals	2	1.0-13	0.01-3.6	3-20	2.4-17	0-0.04	2.0-15	0.01-0.4	2.0-24	1.0-6.0	0.08-7.2	3.0-20	0-1.2	2.0-20
Manufacturing processes														
-Metals	25	0.25-1.5	0.5-1.8	15-58	10-38	0-0.75	2.5-20	0.5-5.0	0.2-7.5	2.5-22	2.8-15	0-5.0	0-0.75	25-138
-Chemicals	5	0.6-7.0	0.1-2.5	2.5-24	1.0-18	0.02-1.5	2.0-15	0-3.0	1.0-6.0	0.4-3.0	0.1-0.4	0.02-2.5	0-0.35	0.2-5.0
-Pulp and paper	3	0.36-4.2	—	0.01-1.5	0.03-0.39	—	0.03-1.5	—	0-0.12	0.01-0.9	0-0.27	0.01-0.9	—	0.09-1.5
-Petroleum products	0.3	0-0.06	—	0-0.21	0-0.06	0-0.02	—	—	0-0.06	0-0.12	0-0.03	0-0.09	—	0-0.24
Atmospheric fallout‡		3.6-7.7	0.9-3.6	2.2-16	6.0-15	0.22-1.8	3.2-20	0.2-1.7	4.6-16	8.7-113	0.44-1.7	0.54-1.1	1.4-9.1	21-58
Dumping of sewage sludge§	[6 × 10 ⁹ kg]	0.4-6.7	0.08-1.3	5.8-32	2.9-22	0.01-0.31	32-106	0.98-4.8	1.3-20	2.9-16	0.18-2.9	0.26-3.8	0.72-4.3	2.6-31
Total input, water		12-70	2.1-17	45-239	35-90	0.3-8.8	109-414	1.8-21	33-194	97-180	3.9-33	10-72	2.1-21	77-375
Median value		41	9.4	142	112	4.6	262	11	113	138	18	41	12	226

* The discharges given represent contaminated process waters, and do not include cooling waters.

† The wastewater production figure corresponds to about 60 m³ capita⁻¹ yr⁻¹ multiplied by the 2.4 × 10⁹ residents in urban and rural areas of the world. The other discharge figures likewise have been derived from the reported water demand per unit tonne of metal smelted or goods manufactured.

‡ We have assumed that 70% of each metal emitted to the atmosphere is deposited on land and the remaining 30% in the aquatic environments^{23,25}.

§ Worldwide sewage sludge production is estimated to be 30 million tonnes, assuming average sludge production rate of 30 g capita⁻¹ day⁻¹ in urban and rural communities^{8,18}. It is believed that 20% of the municipal sludge is directly discharged or dumped into aquatic ecosystems, about 10% is incinerated and the rest is deposited on land.

Using a mass balance receptor model, Pacyna *et al.*³⁰⁻³² obtained a good fit between the atmospheric concentrations at Ny Alesund, Spitsbergen, and the estimated emissions of trace metals for European sources. Such an agreement between the emission estimates and the trace metal distribution in the environment suggests that the model used in the source inventory does yield data that are of the right order of magnitude.

Atmospheric emissions

The emission factors and the statistics on global production or consumption of industrial goods have been used to calculate the worldwide emissions of trace metals to the atmosphere (Table 2). (A blank in this table implies an insignificant contribution from a particular source.) For most of the trace elements, the calculated total emissions vary by a factor of 2-3; the median

values are also shown and may be considered the 'average' rates of global emissions.

Combustion of hard coal, lignites and brown coal in electric power plants and in industrial, commercial and residential burners is the major source of airborne Hg, Mo and Se and a very significant source for As, Cr, Mn, Sb and Tl. Combustion of oil for the same purpose is the most important source of V and Ni and is an important contributor of Sn. The non-ferrous metal industry accounts for the largest fraction of Pb (in addition to gasoline combustion), As, Cd, Cu and Zn emitted. Chromium and Mn are derived primarily from the iron and steel industry (see Table 2).

Our data generally differ from the emission rates that have been reported in the literature. Some of the previous studies have not adequately assessed the important emission sources

Table 5 Worldwide emissions of trace metals into soils (10^6 kg yr⁻¹)*

Source category	Annual global discharge (×10 ¹² kg)	As	Cd	Cr	Cu	Hg	Mn	Mo	Ni	Pb	Sb	Se	V	Zn
Agric. and food wastes	15†	0-6.0	0-3.0	4.5-90	3-38	0-1.5	15-112	9-30	6-45	1.5-27	0-9	0-7.5	3-22	12-150
Animal wastes, manure	2‡	1.2-4.4	0.2-1.2	10-60	14-80	0-0.2	50-140	4-24	3-36	3.2-20	0-0.8	0.4-1.4	2-11	150-320
Logging & other wood wastes	11†	0-3.3	0-2.2	2.2-18	3.3-52	0-2.2	18-104	0-3.3	2.2-23	6.6-8.2	0-5.5	0-3.3	1.1-9.9	13-65
Urban refuse	440§	0.09-0.7	0.88-7.5	6.6-33	13-40	0-0.26	7.0-42	0.22-4.4	2.2-10	18-62	0.22-1.3	0.04-0.62	0-0.4	22-97
Municipal sewage sludge	20	0.01-0.24	0.02-0.34	1.4-11	4.9-21	0.01-0.8	4.4-11	0.08-0.32	5.0-22	2.8-9.7	0.04-0.2	0.01-0.14	0.22-1.5	18-57
Miscellaneous organic wastes including excreta	210¶	0-0.25	0-0.01	0-0.1-0.48	0.04-0.61	—	0.08-0.63	0.06-0.4	0.17-3.2	0.02-1.6	0-0.11	0-0.08	0.11-0.76	0.13-2.1
Solid wastes, metal mfg.	380#	0.01-0.21	0-0.08	0.65-2.4	0.95-7.6	0-0.08	0.41-4.9	0-0.16	0.84-2.5	4.1-11	0-0.16	0-0.19	0.03-0.22	112-484
Coal fly ash and bottom fly ash	3,720**	6.7-37	1.5-13	149-446	93-335	0.37-4.8	498-1,655	15-74	56-279	45-242	2.6-22	4.1-60	11-67	—
Fertilizer	166	0-0.02	0.03-0.25	0.03-0.38	0.05-0.58	—	0.13-0.83	0-0.02	0.20-0.55	0.42-2.3	—	0.02-0.10	0.03-0.13	0.26-1.1
Peat (agricultural and fuel uses)	375††	0.04-0.5	0-0.11	0.04-0.19	0.15-2.0	0-0.02	5.2-17	0.15-0.75	0.22-3.5	0.45-2.6	0.04-0.45	0-0.41	0.08-1.7	0.15-3.5
Wastage of commercial products‡‡		36-41	0.78-1.6	305-610	395-790	0.55-0.82	100-500	0.65-3.2	6.5-32	195-390	0.8-4.0	0.1-0.2	0.6-2.8	310-620
Atmospheric fallout§§ (k)		8.4-18	2.2-8.4	5.1-38	14-36	0.63-4.3	7.4-46	0.55-4.0	11-37	202-263	1.0-3.9	1.3-2.6	3.2-21	49-135
Total input, soils		52-112	5.6-38	484-1,309	541-1,367	1.6-15	706-2,633	30-145	106-544	479-1,113	4.7-47	6.0-76	43-222	689-2,054
Median value		82	22	896	954	8.3	1,670	88	325	796	26	41	132	1,372
Mine tailings (1)		7.2-11	2.7-4.1	—	262-787	0.55-2.8	—	2.1-16	22-64	130-390	16-24	0.28-0.41	1.9-14	194-620
Smelter slags and wastes¶¶ (m)		4.5-9.0	1.6-3.2	—	395-790	0.05-0.28	—	3.2-6.5	32-65	195-390	8-16	0.1-0.2	2.8-6.0	310-620
Total discharge on land		64-132	9.9-45	—	1,198-2,944	2.2-18	—	35-168	160-673	808-1,893	29-87	6.4-77	48-242	1,193-3,294

* The emission figures are derived from the global waste discharges and the estimated emission factors given in Table 2.

† Derived from ref. 45.

‡ This corresponds roughly to the quantity of plants eaten by domestic animals⁴⁵.

§ It is assumed that each urban resident generates 660 g refuse per day¹⁸ and that 25% of all the urban refuse is incinerated.

|| Worldwide sewage sludge production is estimated to be 30 million tonnes, assuming average sludge production rate of 30 g per capita⁻¹ per day in urban and rural communities^{8,18}. It is believed that 20% of the municipal sludge is directly discharged or dumped into aquatic ecosystems, about 10% is incinerated and the rest (70%) is disposed of on land.

¶ This figure represents waste production in non-urban areas equivalent to about 25 g capita⁻¹ day⁻¹.

Assuming that each kg of metal processes or fabricated yields 0.5 kg of slag and waste.

** We have assumed that 75% of each of the trace metals present in coal¹⁰⁻¹² is retained in the 'ashes'. The annual coal production figure is from *Encyclopedia Britannica Ann. Suppl.*, 1987.

†† About 305 × 10⁶ tonnes of peat are used for agriculture and 69 × 10⁶ tonnes used for fuel. We have assumed that 75% of the trace metals in the peat burned are retained in the ashes.

‡‡ We have proposed that 1-15% of the total annual production of the metals may be discarded (lost due to corrosion for instance), or dispersed in soils from usage as chemical, pesticides, crop preservatives, etc. (see text). The production figures used (in million tonnes) are: Cd = 0.014; As₂O₃ = 0.06; Cr(chromite) = 8.9; Sb = 0.08; Co = 0.024; Cu = 7.9; raw steel = 710; Pb = 3.9; Mn = 9.6; Hg = 0.0055; Mo = 0.063; Ni = 0.65; phosphate rock = 137; Se = 0.0014; V = 0.058; Zn = 6.2 (ref. 48).

§§ From Table 2, and the assumption specified in footnote (†) of the same table.

||| We have assumed the average tenor for Pb, Cu, Ni and Zn ore to be 2-6% and that the tailings contain 0.2% of each of these elements. For Mo and V, the ore tenor is assumed to be 0.2-1.5% and the tailings are left with 0.05% of each element. Typical ore tenor for Hg is 0.1-0.5%, and the tailings are assumed to contain 0.02% Hg. For Cd, As, Sb and Se obtained primarily as by-products associated with base metals, it is assumed that 20-30% of each metal content of the ores are left in the tailings.

¶¶ The retention in smelter slags is estimated to be 1-5% of the Hg produced, 5-10% of the Pb, Cu, Ni, Zn, Mo and V, and 10-20% of the As, Cd, Sb and Se produced.

and the emission factors used have not always considered the differences in industrial processes and pollution control strategies. For example, the emission rates reported by Lantzy and Mackenzie²¹ are very different from our estimates. They calculated the industrial emissions on the basis of published chemical composition of metal-enriched, fine-grained aerosols, and estimated the fossil fuel contributions assuming, quite erroneously, that 90% of the metal concentrations in the oil and coal is released to the atmosphere. Their emission rate for lead also excluded the very important automotive contribution.

Previous estimates^{23,24,33} of anthropogenic emissions of Hg generally fall in the range 2,000-10,000 tonnes, and are in agreement with our own inventory. The reported As emission rates of 23,600-28,000 tonnes^{20,25} fall close to the maximum values in Table 2. The total (volatile + particulate) Se emission of 6,700-8,300 tonnes recently reported by Ross²⁶ is in good agreement with our own estimate of 3,020-9,625; it is gratifying to note that the Se emission rate of 6,000 tonnes yr⁻¹ recently reported by Mosher and Duce³⁴ is identical to our own estimate (see footnote, Table 2). The inventories for 1975 previously reported by Nriagu² were 56,000, 449,000, 470,000 and 314,000 tonnes respectively for Cu, Pb, Ni and Zn. The current (1983/84) emission rates for Cu, Pb and Zn are lower than those of 1975, due mostly to the overall global reduction in the emission of particulates by industries and the phase down in the consumption of leaded gasoline³⁵. It seems that the rate for Ni in 1975 was underestimated.

Mean global emissions rates from natural sources have recently been estimated to be (in tonnes per year) 7,800 for As, 1,000 for Cd, 5,400 for Co, 19,000 for Cu, 516,000 for Mn, 26,000 for Ni, 19,000 for Pb, 66,000 for V, 46,000 for Zn (all from ref. 3) and 6,000 for Hg (ref. 24). The most recent estimate suggests that 6,000-13,000 tonnes of Se are annually released to the

atmosphere from natural sources with 60-80% of the total Se emission being of marine biogenic origin³⁴. A comparison of these fluxes from natural sources with the anthropogenic emissions in Table 2 leaves no doubt as to the influence of industrial activities on the atmospheric cycle of the trace elements. On average the anthropogenic emissions of As, Cd, Cu, Ni and Zn exceed the inputs of these elements from natural sources by about two-fold or more; in the case of lead, the ratio of anthropogenic to natural emission rates is about 17.

Discharges into water

The main industrial use for water is in the cooling system. We have, however, considered only the generation of contaminated process waters in deriving the inventories in Table 4. The major sources of trace metal pollution in aquatic ecosystems including the ocean are domestic wastewater effluents (especially As, Cr, Cu, Mn and Ni), coal-burning power plants (As, Hg and Se in particular), non-ferrous metals smelters (Cd, Ni, Pb and Se), iron and steel plants (Cr, Mo, Sb and Zn) and the dumping of sewage sludge (As, Mn and Pb). The atmosphere is the major route of Pb entry in natural waters, a fact that has been well documented in the literature³⁶⁻³⁸. The atmosphere also accounts for over 40% of the V loading, a surprising observation in so far as little is currently known about the atmospheric chemistry of this element.

We are not aware of any published inventory pertaining to the global discharges of metal pollution into the aquatic ecosystems; excellent studies on the marine cycle of trace metals are available however³⁶⁻⁴⁰. We infer from Table 4 that, for most of the trace metals, the annual anthropogenic inputs into the water exceed the quantities emitted to the atmosphere. Although air pollution by toxic metals has been recognized as a matter of concern, the impacts of loading large quantities of toxic metals

into the freshwater resources remain to be fully assessed on the global or regional scale.

If it is assumed that only 25% of the industrial effluents are discharged into lakes and rivers (total volume, 1.3×10^{16} l), the average concentrations in these waters (based on 25% of the median values in Table 4) would be increased by about 90 ng l^{-1} for Hg, 180 ng l^{-1} for Cd, 800 ng l^{-1} for Se and As, $2,200 \text{ ng l}^{-1}$ for Cu and Ni, $\sim 2,500 \text{ ng l}^{-1}$ for Zn and over $4,000 \text{ ng l}^{-1}$ for Pb. The background concentrations of trace metals in unpolluted lakes and rivers (see refs 41–43 for example) are, in general, several-fold lower than these expected increases. In other words, the current rate of worldwide industrial inputs greatly exceed the baseline burdens of trace metals in the average lake and river. Most of the effluent discharges occur in Europe, North America and some Asian countries, implying that the contamination of the freshwater resources in these regions may be much more severe than is generally realized. This problem has not elicited much discussion because (1) the available data bases are often inadequate for assessing the degree of metal contamination of many lakes and rivers, and (2) the short half lives of trace metals (due to their rapid transfer to the sediments) tend to reduce the concentrations of pollutant metals in the water column⁴⁴.

Discharges into the soil

Our inventory (Table 5) clearly suggests that soils are receiving large quantities of trace metals from a wide variety of industrial wastes. The two principal sources of trace metals in soils, however, are the disposal of ash residues from coal combustion and the general wastage of commercial products on land. Urban refuse represents an important source of Cu, Hg, Pb and Zn with notable contributions of Cd, Pb and V also coming via the atmosphere. The large volumes of wastes associated with animal husbandry, logging as well as agricultural and food production can affect the trace metal budget of many soils significantly (Table 5). Although municipal sewage sludge may not be a particularly important source on a global scale, its trace metal

content is often so high that it is sometimes unsuitable for disposal on land. On a local scale, municipal sewage represents one of the most important sources of metal contamination in soils.

If the total metal inputs were dispersed uniformly over the cultivated land area of $16 \times 10^{12} \text{ m}^2$ (ref. 45), the annual rates of metal application would vary from about 1.0 g ha^{-1} for Cd and Sb to about 50 g ha^{-1} for Pb, Cu and Cr to over 65 g ha^{-1} for Zn and Mn. Although discernable increases have been noted in the metal burdens of some surface soils^{29,46}, the large background reservoir of trace metals generally obscures such huge loadings from industrial sources. Nevertheless, each soil has a limited retention capacity for trace metals and there is growing concern that many soils in Japan and central Europe either have become or will soon become overloaded with toxic metals at the current rate of anthropogenic input^{29,46,47}. The technology for decontaminating such soils has yet to be developed.

Conclusions

The inventories presented here clearly show that mankind has become the most important element in the global biogeochemical cycling of the trace metals. The man-induced mobilization of trace metals into the biosphere (median values in thousand tonnes yr^{-1} of the terrestrial plus aquatic inputs minus atmospheric emissions) comes to about 120 for As, 30 for Cd, 2,150 for Cu, 11 for Hg, 110 for Mo, 470 for Ni, 1,160 for Pb, 72 for Sb, 79 for Se, 71 for V, and 2,340 for Zn. The annual total toxicity of all the metals mobilized, in fact, exceeds the combined total toxicity of all the radioactive and organic wastes generated each year, as measured by the quantity of water needed to dilute such wastes to drinking water standard. Each year, millions of tonnes of 'new' trace metals are produced from the mines and subsequently redistributed in the biosphere. The greatly increased circulation of toxic metals through the soils, water and air and their inevitable transfer to the human food chain remains an important environmental issue which entails some unknown health risks for future generations.

Received 1 October 1987; accepted 1 March 1988.

1. *Changing Metals Cycles and Human Health* (ed. Nriagu, J. O.) (Springer, Berlin, 1984).
2. Nriagu, J. O. *Nature* **279**, 409–411 (1979).
3. Pacyna, J. M. *Adv. Environ. Sci. Technol.* **17**, 1–32 (1986).
4. Pacyna, J. M. in *Lead, Mercury, Cadmium and Arsenic in the Environment* (eds Hutchinson, T. C. & Meema, K. M.) 69–87 (Wiley, New York, 1987).
5. Adriano, D. C. *Trace Elements in the Terrestrial Environment* (Springer, New York, 1986).
6. *Minerals and the Environment* (ed. Jones, M. J.) (Institution of Mining and Metallurgy, London, 1974).
7. Williamson, N. A., Johnson, M. S. & Bradshaw, A. D. *Mine Waste Reclamation* (Mineral Industry Research Organization, London, 1982).
8. Horvath, D. J. & Koshut, R. A. *J. Environ. Qual.* **10**, 491–495 (1981).
9. Campbell, W. J. *Environ. Sci. Technol.* **10**, 436–439 (1976).
10. Hasanen, E., Pohjola, V., Hahkala, M., Zilliacus, R. & Wickstrom, K. *Sci. Total Environ.* **54**, 29–51 (1986).
11. Mumma, R. O. *et al. Arch. Environ. Contam. Toxicol.* **13**, 75–83 (1984).
12. Maninini, G. & Santori, M. *Agric. Ecosystems Environ.* **18**, 291–311 (1987).
13. Kuusela, K. *Ambio* **16**, 80–85 (1987).
14. Dyer, J. C. & Mignone, N. A. *Handbook of Industrial Residues* (Noyes Publications, Park Ridge, NJ, 1983).
15. Sittig, M. *Metal and Inorganic Waste Reclaiming Encyclopedia* (Noyes Data Corporation, Park Ridge, NJ, 1980).
16. Kugelman, I. J. *Toxic and Hazardous Wastes* (Lehigh University Press, 1985).
17. Shuler, M. L. *Utilization and Recycle of Agricultural Wastes and Residues* (CRC Press, Boca Raton, FL, 1980).
18. Bond, R. G. & Straub, C. P. *Handbook of Environmental Control, Vol. II: Solid Wastes* (CRC Press, Cleveland, 1973).
19. Culp, G., Wesner, G., Williams, R. & Hughes, M. V. *Wastewater Reuse and Recycling Technology* (Noyes Data Corporation, Park Ridge, NJ, 1980).
20. Walsh, P. R., Duce, R. A. & Fasching, J. L. *J. geophys. Res.* **84**, 1719–1726 (1979).
21. Lantzy, R. J. & Mackenzie, F. T. *Geochim. cosmochim. Acta* **43**, 511–523 (1979).
22. Jaworowski, Z., Bysiek, M. & Kownacka, L. *Geochim. cosmochim. Acta* **45**, 2185–2199 (1981).
23. Mackenzie, F. T. & Wollast, R. in *The Sea Vol. 6* (eds Goldberg, E. D., McCave, I. N., O'Brien, J. J. & Steele, J. H.) 739–785 (Wiley, New York, 1977).
24. Lindquist, O. & Rodhe, H. *Tellus* **37B**, 136–159 (1985).
25. Chilvers, D. C. & Peterson, P. J. in *Lead, Mercury, Cadmium and Arsenic in the Environment* (eds Hutchinson, T. C. & Meema, K. M.) 279–301 (Wiley, New York, 1987).
26. Ross, H. B. *Tellus* **37B**, 78–90 (1985).
27. Boutron, C. F. *Adv. Environ. Sci. Technol.* **17**, 446–505 (1986).
28. *Historical Monitoring*, MARC Report 31 (Chelsea College, University of London, 1985).
29. Jones, K. C., Symon, C. J. & Johnston, A. E. *Sci. Total Environ.* **61**, 131–144 (1987).
30. Pacyna, J. M., Ottar, B., Tomza, U. & Maenhaut, W. *Atmos. Environ.* **19**, 857–865 (1985).
31. Pacyna, J. M. *Wat. Air Soil Pollut.* **30**, 825–835 (1985).
32. Pacyna, J. M. *Atmos. Environ.* **18**, 41–50 (1984).
33. Watson, W. D. in *The Biogeochemistry of Mercury in the Environment* (ed. Nriagu, J. O.) 41–77 (Elsevier, Amsterdam, 1979).
34. Mosher, B. W. & Duce, R. A. *J. Geophys. Res.* **92**, 13289–13298 (1987).
35. *State of the Environment: A View Toward the Nineties* (Conservation Foundation Washington, DC, 1987).
36. Flegal, A. R. & Patterson, C. C. *Earth planet. Sci. Lett.* **64**, 19–32 (1983).
37. Boyle, E. A., Chapnick, S. D. & Shen, G. T. *J. geophys. Res.* **91**, 8573–8593 (1986).
38. Veron, A., Lambert, C. E., Isley, A., Linet, P. & Grousset, F. *Nature* **326**, 278–281 (1987).
39. *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) (Plenum, New York, 1983).
40. *The Role of Air-Sea Exchange in Geochemical Cycling* (ed. Baut-Menard, P.) (Reidel, Dordrecht, 1985).
41. Nriagu, J. O. *Sci. Total Environ.* **58**, 89–92 (1986).
42. Grant, B., Ming-hui, H., Boyle, E. A. & Edmond, J. M. *EOS Trans. Am. Geophys. Union* **63**, 48 (1982).
43. Martin, J. M. & Whitfield, M. in *Trace Metals in Sea Water* (eds Wong, C. S. *et al.*) 265–296 (Plenum, New York, 1983).
44. Nriagu, J. O. in *The Role of the Oceans as a Waste Disposal Option* (ed. Kullenberg, G.) 441–468 (Reidel, Dordrecht, 1986).
45. Diamond, J. M. *Nature* **328**, 479–480 (1987).
46. Asami, T. in *Changing Metal Cycles and Human Health* (ed. Nriagu, J. O.) 95–111 (Springer, Berlin, 1984).
47. Kloke, A., Sauerbeck, D. R. & Vetter, J. in *Changing Metal Cycles and Human Health* (ed. Nriagu, J. O.) 113–141 (Springer, Berlin, 1984).
48. *Canadian Minerals Yearbook 1985* (Energy, Mines and Resources Canada, Canadian Government Publishing Centre, Ottawa, 1985).

A simple structural feature is a major determinant of the identity of a transfer RNA

Ya-Ming Hou & Paul Schimmel

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Analysis of a series of mutants of an Escherichia coli alanine transfer RNA shows that substitution of a single G-U base pair in the acceptor helix eliminates aminoacylation with alanine in vivo and in vitro. Introduction of that base pair into the analogous position of a cysteine and a phenylalanine transfer RNA confers upon each the ability to be aminoacylated with alanine. Thus, as little as a single base pair can direct an amino acid to a specific transfer RNA.

THE elucidation of the genetic code revealed the essential role played by transfer RNAs in decoding genetic information. The members of this class of molecules each have an identity defined by the specific amino acid attached to the 3' end. This attachment is carried out by an aminoacyl tRNA synthetase. The molecular determinants which define the identity of a tRNA are unknown, but these determinants are presumably sites that are specifically recognized by aminoacyl tRNA synthetases.

There is evidence that, in at least some synthetase-tRNA complexes, the bound protein makes contact with the tRNA anticodon¹⁻³. While this sequence distinguishes the various tRNAs from each other, and is important for synthetase recognition in some instances^{4,5}, it is not the general basis upon which the identity of each tRNA is established. This is clear from the behaviour of some amber suppressor tRNAs (CUA anticodon) which preserve their identity in spite of a change in anticodon sequence. The variability in behaviour amongst tRNAs suggests that the locations of the determinants for tRNA identity are idiosyncratic.

Biochemical and physical methods have been used to investigate the structural features of synthetase-tRNA complexes, and tRNA sequence variants have also been investigated as substrates *in vivo* and *in vitro*^{1-3,6}. Although a specific sequence alteration in a tRNA can lead to a reduction or loss of aminoacylation capacity, this does not establish that the site of alteration is a determinant of the identity. Similarly, sites on tRNAs that crosslink to bound aminoacyl tRNA synthetases are not necessarily sites that are crucial for recognition. This is because parts of a bound protein that are close to the nucleic acid are not necessarily making base- and amino-acid-specific contacts.

For these reasons, the determinants for the identity of a tRNA can only be established by defining the minimal set of nucleotides that is sufficient to confer amino-acid specificity. Ideally, it should be possible to transfer this set of nucleotides from one framework to another and to show that identity is transferred with that set of bases. Such manipulations must be done within the constraints of a group of conserved nucleotides which have to be preserved in order to maintain highly specific tertiary interactions. With these constraints in mind, Normanly *et al.* recently presented evidence that no more than 12 nucleotides are required to convert a leucine- to a serine-specific tRNA⁷.

A series of studies of *E. coli* alanyl tRNA synthetase have led to the delineation of a segment of the polypeptide that is critical for binding to tRNA^{Ala} (refs 8-10). To further investigate the tRNA-protein complex, we sought to obtain mutants in the tRNA that were defective in recognition by this enzyme. These mutants could then be used to obtain and map second-site revertants in the protein that compensate for the defect in the mutant tRNA. As shown below, we found that a large number of sequence alterations can be introduced into tRNA^{Ala} without loss of recognition by the alanine enzyme. This suggested that a simple structural feature is a major determinant for establishing

the identity of this tRNA. Further analysis showed that this structural feature is a unique base pair in the amino-acid acceptor helix, and that introduction of this base pair into two other tRNAs enables each to be recognized by the alanine enzyme.

Experimental design

Genes for three amber-suppressing tRNAs were used as starting materials for all of the mutants generated. These genes encode tRNA^{Cys/CUA}, tRNA^{Phe/CUA} and tRNA^{Ala/CUA}, and are each harboured on the plasmid vector pGF1B-1 (refs 11, 12). Transcription of the tRNA genes in pGF1B-1 is initiated at the strong *Ipp* promoter¹³. Each of the three tRNA genes produces a functional suppressor which retains the respective amino-acid specificity, in spite of the change to the CUA anticodon. The sequence of the gene for tRNA^{Ala/CUA} is based on that of tRNA^{Ala/GGC} (ref. 14), except for the substitution of the anticodon triplet and a change of U38 → A to improve the efficiency of amber suppression¹⁵.

A modified oligonucleotide-directed mutagenesis procedure was used to generate a population of mutants in this gene. These mutants included single- and multiple-base substitutions, and were designed so that they would preserve the tertiary interactions in the molecule. The objective was to generate one or more mutants defective in aminoacylation by alanine tRNA synthetase. Because of evidence that synthetases interact with the 'inside' of the L-shaped tRNA structure¹⁶, we directed many mutations to this part of the molecule. The fourth position from the 3' end of the molecule was also varied because this position has been proposed as a 'discriminator' base which enables synthetases to distinguish between classes of tRNA molecules¹⁷.

To test for recognition of mutant tRNAs by the alanine enzyme, we needed an *E. coli* host strain that harbours an amber-containing gene, the phenotype of which could only be suppressed by insertion of alanine. With such a strain, no sup⁺ phenotypes would arise from misaminoacylation of a mutant tRNA^{Ala/CUA} with an amino acid other than alanine. The closest approximation we obtained was *E. coli* strain FTP3689 (*trpA*(UAG234))¹⁸ which has a deletion of the *trpAB* region and harbours an F' episome that encodes *trpA*(UAG234). This *trpA* allele contains an amber mutation at position 234 of the alpha subunit of tryptophan synthetase, so that a strain carrying it is a tryptophan auxotroph. Only insertion of glycine or alanine are known to suppress the Trp⁻ phenotype of *trpA*(UAG234)¹⁸. Therefore, except for the possibility of a sup⁺ phenotype arising from misacylation of mutant tRNA^{Ala/CUA} with glycine, strain FTP3689 is well suited for these studies.

Suppression activity of mutant tRNAs

Figure 1a shows the sequence and cloverleaf structure of 'wild-type' tRNA^{Ala/CUA} which, as stated above, differs in four positions from the sequence of tRNA^{Ala/GGC}. Including these four changes, a total of 36 substitutions have been made collectively

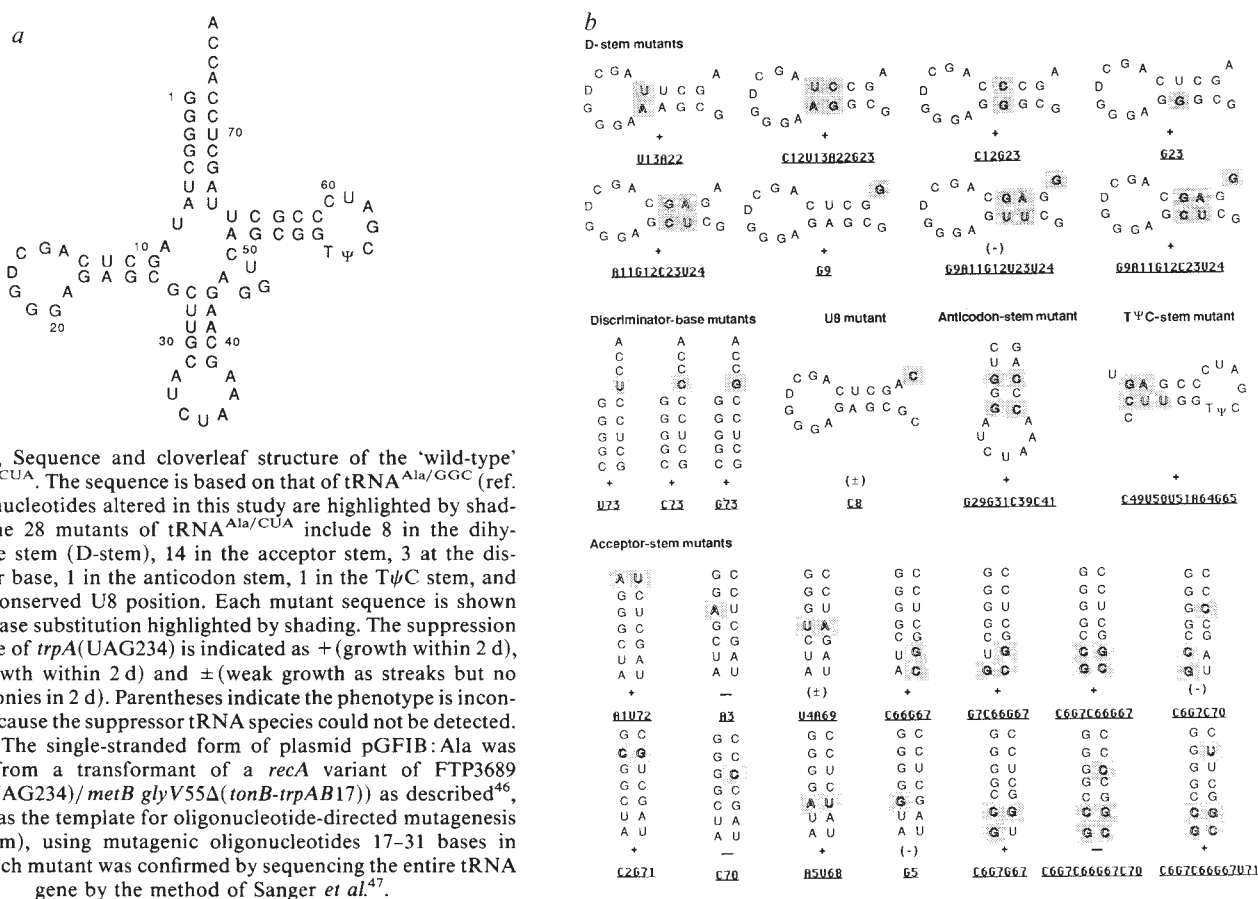


Fig. 1 *a*, Sequence and cloverleaf structure of the 'wild-type' tRNA^{Ala/CUA}. The sequence is based on that of tRNA^{Ala/GGC} (ref. 14). The nucleotides altered in this study are highlighted by shading. *b*, The 28 mutants of tRNA^{Ala/CUA} include 8 in the dihydrouridine stem (D-stem), 14 in the acceptor stem, 3 at the discriminator base, 1 in the anticodon stem, 1 in the TψC stem, and 1 at the conserved U8 position. Each mutant sequence is shown with the base substitution highlighted by shading. The suppression phenotype of *trpA*(UAG234) is indicated as + (growth within 2 d), - (no growth within 2 d) and ± (weak growth as streaks but no single colonies in 2 d). Parentheses indicate the phenotype is inconclusive because the suppressor tRNA species could not be detected. **Methods.** The single-stranded form of plasmid pGFIB:Ala was purified from a transformant of a *recA* variant of FTP3689 (*F' trpA*(UAG234)/*metB glyV55Δ*(*tonB-trpAB17*)) as described⁴⁶, and used as the template for oligonucleotide-directed mutagenesis (Amersham), using mutagenic oligonucleotides 17–31 bases in length. Each mutant was confirmed by sequencing the entire tRNA gene by the method of Sanger *et al.*⁴⁷.

among all of the mutant tRNA^{Ala/CUA}s and the positions of these changes are indicated by shading on the tRNA cloverleaf. All but one of these substitutions (U8 → C) are at non-conserved positions so that, collectively, over one-half of the 61 variable positions have been changed. The 28 mutant tRNA^{Ala/CUA} species that have been investigated are shown in Fig. 1*b*. The suppressor phenotype (+ or -) as tested on the *trpA*(UAG234) allele is indicated for each; the parenthetical (-) designation indicates that the particular suppressor tRNA species could not be detected, so that the '-' suppressor phenotype is inconclusive. Many of the 36 substitutions lie along the inside of the L-shaped three-dimensional tRNA structure.

Eight mutants with one or more alterations in the dihydrouridine stem were created. One of these has five substitutions (G9A11G12C23U24) that recreates the sequence of the dihydrouridine stem of tRNA^{Met}. This, and six of the other mutants in this group have a sup⁺ phenotype, while the phenotype of the remaining mutant is inconclusive. These results are intriguing because considerable data indicate that synthetases come into close contact with the dihydrouridine stem¹.

It is worth noting that the mutant anticodon stem shown in Fig. 1*b* creates a sequence that has a feature common to all prokaryotic and eukaryotic initiator tRNAs (namely, three consecutive G-C base pairs¹⁹) and that the mutant TψC stem is identical in sequence to that of yeast tRNA^{Phe}. In spite of these base substitutions, each mutant has a sup⁺ phenotype. Similarly, each of the three discriminator-base mutants (position 73) is active in suppression of the *trpA*(UAG234) allele.

Altogether 20 of the 28 mutants suppressed the *trpA*(UAG234) allele, while results with five of the mutants are inconclusive because the mutant tRNA species could not be detected by our screen (see below). Only three mutants (A3, C70, and C667C66G67C70) have a clear sup⁻ phenotype. These mutants share in common a change in the G3:U70 base pair that is

located in the acceptor helix.

We estimated the suppression efficiency of three of the sup⁺ mutant tRNAs which had multiple nucleotide substitutions: the anticodon-stem mutant (G29G31C39C41), a dihydrouridine-stem mutant (G9A11G12C23U24) and the TψC-stem mutant (C49U50U51A64G65). We confirmed that, in addition to suppression of the *trpA*(UAG234) amber codon, the wild-type tRNA^{Ala/CUA} and each of these variants also suppress the *trpA*(UAG15) amber allele²⁰. This established that each mutant tRNA^{Ala/CUA} can suppress amber mutations in at least two different regions of the *trpA* messenger RNA. The suppression efficiencies at both amber codons were measured by determining the activity of the beta-subunit of tryptophan synthetase, in the presence and absence of the alpha-subunit²¹. (Association of the beta-subunit with the alpha-subunit greatly enhances the activity of the former, so that the extent of enhancement provides a quantitative measure of the amount of *trpA*-encoded alpha-subunit present²².) These assays showed that the suppression efficiency at the *trpA*(UAG234) and *trpA*(UAG15) alleles is the same (21% and 20%, respectively) for the 'wild-type' tRNA^{Ala/CUA} suppressor. For the mutants, the suppression efficiencies at the UAG234 and UAG15 amber alleles, respectively are: anticodon-stem mutant (21%, 46%); dihydrouridine-stem mutant (12%, 7%); and TψC-stem mutant (5%, 18%). We also determined a suppression efficiency of 12% for the discriminator-base U73 mutant at the UAG15 allele. These measurements show that the mutant tRNAs have suppression efficiencies which, depending on the allele, are at least 25% of the efficiency of the wild-type suppressor.

In the *trpA*(UAG15) system, where insertion of a variety of amino acids leads to suppression²⁰, we detected weak suppression with the three acceptor stem mutants (A3, C70, and C667C66G67C70) that do not suppress the *trpA*(UAG234) allele. The suppression efficiencies, however (<3%) are near

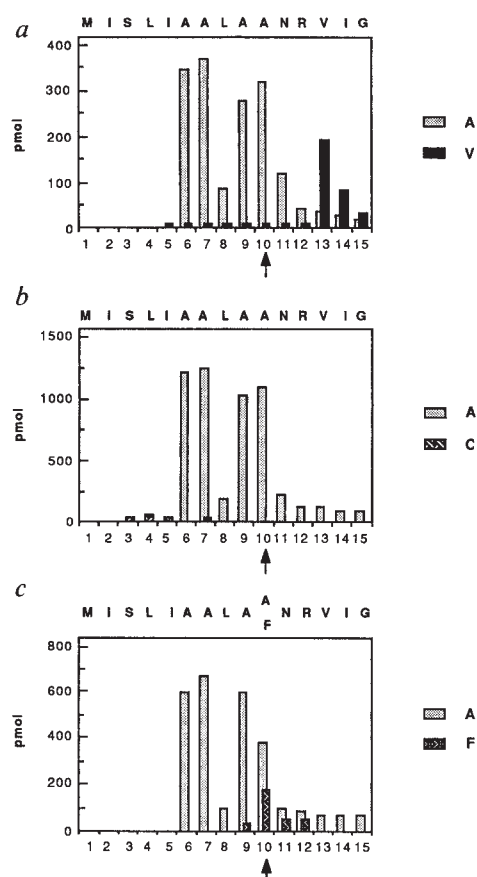


Fig. 2 Amino-acid sequence analysis of the first 15 residues of DHFR. The picomolar yields of the indicated amino acids (alanine, valine, phenylalanine, or cysteine) are plotted versus the residue number for the *fol* gene products whose amber codon at position 10 was suppressed by the G9A11G12C23U24 tRNA^{Ala/CUA} sup⁺ mutant (a), the G3:U70 tRNA^{Cys/CUA} mutant (b) and the G3:U70 tRNA^{Phe/CUA} mutant (c). The sequence of the gene product is given across the top of each diagram.

Methods. Wild-type DHFR was purified from *E. coli* strain X9100.1 (F'*lacI*^qZ::tn5YA/ Δ (*lac proB*) *nalA rif argC_{am} ara*) (kindly provided by the laboratory of J. Abelson) harbouring plasmid pTAC101, as described⁷. Variants of DHFR were purified separately from X9100.1 strains carrying the tRNA suppressor gene on plasmid pGFIB-I and the *fol_{am-10}* gene on a compatible plasmid derived from pACYC184 (ref. 48). The *fol_{am-10}* allele was created by altering⁷ the sequence of the *fol* gene (also provided by the laboratory of J. Abelson) at codons 10 and 11 from those of valine-aspartic acid to amber-asparagine. Carboxymethylation of the cysteine residues in DHFR was accomplished by treating 10 nmol of the denatured protein (in 0.1 M Tris-HCl (pH 7.0), 0.1 mM dithiothreitol (DTT), 6 M guanidine-HCl) with 60 nmol of iodoacetic acid (in 0.1 M Tris-HCl (pH 8.0)) and 300 nmol of [¹⁴C]-iodoacetic acid (18 mCi mmol⁻¹, New England Nuclear) for 1 h in the dark at room temperature. The reaction was terminated by addition of 300 nmol of DTT. Before sequence analysis, the protein was dialysed overnight against water. Protein sequences were determined at the microsequencer facility at the Whitehead Institute (MIT) or at the Microchemical Laboratory at the University of Southern California.

the background level. The implication that these mutants are severely defective for aminoacylation with alanine is supported by further evidence given below.

The amino acid inserted by several of the sup⁺ mutant tRNAs was checked by protein-sequence analysis using the plasmid-encoded mutant *fol* gene (*fol* encodes dihydrofolate reductase (DHFR)) which has an amber codon at position 10 (ref. 7). The protein resulting from suppression with a mutant tRNA^{Ala/CUA}

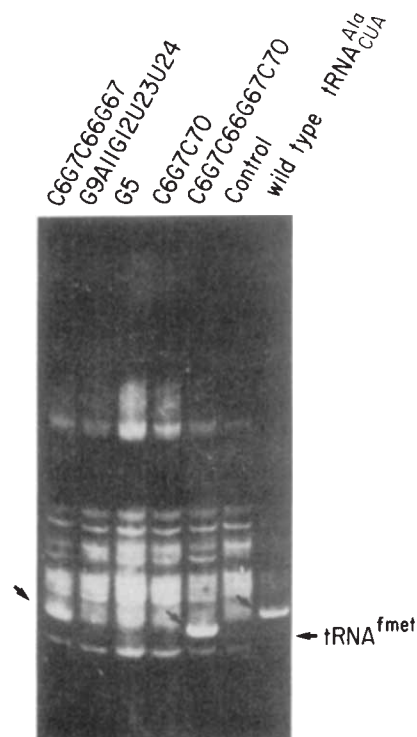


Fig. 3 Synthesis of mutant tRNA^{Ala/CUA} analysed by gel electrophoresis. Total tRNA was purified from a *recA* version of FTP3689 carrying constructs of pGFIB:Ala which encode genes for 'wild-type' and mutant tRNA^{Ala/CUA}. The position of the tRNA^{fMet} species is indicated at the front of the gel²⁴. The positions of those tRNA^{Ala/CUA} species which can be detected are marked with arrows.

Methods. A 50 ml culture of strain FTP3689 was grown to late log phase in L broth⁴⁹. Cells were collected, resuspended in 1/100 volume of 10 mM Tris-HCl (pH 7.4), 1 mM MgCl₂, and extracted for 1 min with water-saturated phenol. After centrifugation for 30 min at 8,000 $\times$ g (4 $^{\circ}$ C), the cells were again subject to phenol extraction. The combined aqueous phase was precipitated with ethanol and the pellet dissolved in 10 mM Tris-HCl (pH 8.0), 1 mM MgCl₂. Approximately 10 μ g of each pellet was loaded in each lane of an analytical 12% polyacrylamide gel (29:1). The gel (20 cm $\times$ 25 cm $\times$ 0.75 cm) was run in a buffer containing 89 mM Tris borate, 89 mM boric acid, 3 mM EDTA (pH 8.0) at 500 V (30 mA) with the temperature kept below 35 $^{\circ}$ C until the xylene cyanole dye had migrated 20 cm. The tRNAs were visualized by staining with ethidium bromide.

species was purified by methotrexate affinity chromatography^{7,23} and sequenced. The following four mutant sup⁺ tRNAs were investigated: discriminator-base (U73); dihydrouridine-stem: (G9A11G12C23U24); anticodon-stem (G29G31C39C41); and T ψ C-stem (C49U50U51A64G65). For each of these, sequence analysis of the suppressed *fol* gene product showed that alanine was present at position 10 greater than 97% of the time.

As an example, Fig. 2a shows the alanine and valine elution profiles for the first 15 positions for the product of the plasmid-encoded *fol* gene which has been suppressed with the dihydrouridine-stem (G9A11G12C23U24) tRNA^{Ala/CUA} sup⁺ mutant. In this case only alanine is detected at position 10, because the suppressed, plasmid-encoded protein is far more abundant than the chromosome-encoded copy which has a valine at position 10. When the *fol* gene was suppressed by the other sup⁺ mutants investigated, we typically detected <3% valine at position 10.

The appearance of alanine at positions 8 and 11, and of valine at position 14, is due to carry-over from the preceding position (position 7, 10, and 13, respectively). Except for carry-overs,

only the expected amino acid was detected at each of the first 15 positions for all sequence analyses performed.

Aminoacylation of mutant tRNAs

The plasmid-encoded mutant tRNAs were visualized by non-denaturing electrophoresis of samples from phenol-treated cells on a 12% polyacrylamide gel (Fig. 3). The observed mobility of the different tRNAs is approximately in accordance with their GC content²⁴, with the GC-rich tRNA^{Met} species migrating the fastest. The 'wild-type' tRNA^{Ala/CUA} species is readily detected in this system because of the large amounts produced from the multicopy plasmid. Similarly, the sup⁺ C6G7C66G67 mutant and sup⁻ C6G7C66G67C70G70 (not shown) and A3 (not shown) mutant tRNAs are easily visualized. On the other hand, the G9A11G12U23U24G5, and C6G7C70 mutants cannot be detected and it is for this reason that no conclusion can be drawn about their suppression phenotypes.

The gel electrophoresis system provides a way to isolate sufficient amounts of the mutant tRNAs for analytical aminoacylation measurements. We investigated five species: 'wild-type' tRNA^{Ala/CUA}; the sup⁺ C6G7C66G67 acceptor-stem mutant; and the three sup⁻ acceptor-stem mutants which have an altered G3:U70 base pair (A3, C70, and C6G7C66G67C70). The purity of the tRNA^{Ala/CUA} species (purified from the gel) was estimated to be at least 90%, as determined by the maximal pmoles of alanine incorporated ($\geq 1,600$) per A₂₆₀ absorbance unit. Each of these tRNAs, at a concentration of 10 μ M, was incubated with highly purified *E. coli* alanine tRNA synthetase. Under the conditions of the assay, the K_m for the 'wild-type' tRNA^{Ala/CUA} was determined to be 1 μ M. This compares with a K_m of 0.65 μ M measured previously for a mixture of tRNA^{Ala/GGC} and tRNA^{Ala/UGC} isoacceptors contained within unfractionated tRNA²⁵. We found that there is only a small difference in the time-course of aminoacylation of the sup⁺ C6G7C66G67 mutant versus the purified tRNA^{Ala/UGC} isoacceptor and 'wild-type' tRNA^{Ala/CUA} (Fig. 4). In contrast, we could detect no aminoacylation by alanine tRNA synthetase of the three sup⁻ species with an altered G3:U70 base pair.

Transfer of the G3:U70 base pair

To investigate whether the G3:U70 base pair is a major determinant for the identity of an alanine-specific tRNA, we introduced this base pair into two other tRNAs which differ substantially in sequence from tRNA^{Ala}. The tRNA^{Cys/CUA} and tRNA^{Phe/CUA} amber suppressors differ from tRNA^{Ala/CUA} at 38 and 31 positions respectively (Fig. 5a). These differences are dispersed throughout the sequences of the two molecules. The regions of greatest sequence divergence from tRNA^{Ala/CUA} are the anticodon stem and dihydrouridine stem and loop of tRNA^{Cys/CUA}, and the anticodon and acceptor stems of tRNA^{Phe/CUA}. While both tRNA^{Ala/CUA} and tRNA^{Phe/CUA} are type 1 tRNAs, which have four base pairs in the dihydrouridine stem and five nucleotides in the variable loop, tRNA^{Cys/CUA} has only three base pairs in the dihydrouridine stem and four in the variable loop.

The unaltered tRNA^{Cys/CUA} and tRNA^{Phe/CUA} suppressors were introduced into *E. coli* strain FTP3689(*trpA*(UAG234)) and the resulting transformants were Trp⁻. This is expected because neither Cys nor Phe at position 234 gives an active alpha-subunit of tryptophan synthetase. Each of these tRNAs has a C3:G70 base pair. This base pair was replaced with a G3:U70 pair and plasmids encoding the respective mutant tRNAs, when introduced into strain FTP3689, conferred a Trp⁺ phenotype (data not shown). This result suggested that installation of the G3:U70 base pair caused these tRNAs to be recognized by alanine tRNA synthetase.

The tRNA^{Cys/CUA} and tRNA^{Phe/CUA} suppressors insert Cys and Phe, respectively¹¹. Figure 2b shows the elution profiles for alanine and cysteine for the *fol* gene product that has been suppressed by the G3:U70 tRNA^{Cys/CUA} sup⁺ mutant. For this

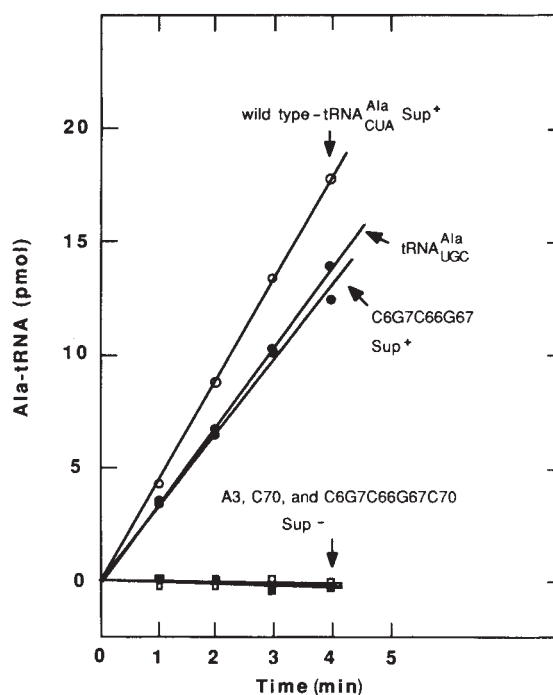


Fig. 4 Initial velocity of aminoacylation activity of the 'wild-type' tRNA^{Ala/CUA} and the C6G7C66G67, A3, C70, and C6G7C66G67C70 mutants. Assays were performed at 23 °C as described³⁵. Purified alanyl-tRNA synthetase (0.065 μ M in 5 μ l) was added to an assay mixture (45 μ l) containing 10 mM MgCl₂, 5 mM DTT, 4 mM ATP, 40 mM sodium phosphate (pH 7.5), 20 μ M alanine, 2.4 μ M [2,3-³H] alanine and 10 μ M tRNA. Each tRNA^{Ala/CUA} species was purified by gel elution from a preparative 12% polyacrylamide gel (20 cm $\times$ 25 cm $\times$ 2.5 cm)^{24,49}. The wild-type tRNA^{Ala/UGC} isoacceptor was purchased from Subriden RNA.

sample the cysteine residues in DHFR were carboxymethylated with ¹⁴C-iodoacetic acid. The data show that introduction of the G3:U70 base pair results in insertion of alanine at position 10. In fact, there is no significant insertion of cysteine by the mutant tRNA and the peak ratio of alanine at position 10 relative to position 9 is identical to the peak ratio between positions 7 and 6. In addition, no other amino acid is detected at position 10. These data suggest virtually complete conversion of the cysteine-specific suppressor to an alanine-specific suppressor through a simple C3:G70 $\rightarrow$ G3:U70 replacement.

Figure 2c shows the results obtained on DHFR that was isolated from the strain harbouring the G3:U70 tRNA^{Phe/CUA} sup⁺ mutant. Both alanine and phenylalanine are present at position 10. We estimate that ~20% of the alanine at this position results from carry-over from position 9. This is based on the amount of alanine present at position 8 that carries over from position 7. When the amount of alanine at position 10 is corrected for the carry-over from position 9, a ratio of 1.7 moles of Ala per mole of Phe is computed. The same ratio of Ala to Phe was obtained at two different microsequenator facilities. The results show that the G3:U70 tRNA^{Phe/CUA} mutant suppressor is recognized *in vivo* by alanine tRNA synthetase.

The tRNA^{Cys/CUA} suppressor and the G3:U70 variant were investigated *in vitro* by analytical aminoacylation measurements. Under the conditions used, the 'wild-type' tRNA^{Cys/CUA} is not a substrate for the purified alanine tRNA synthetase, even with substrate levels of the enzyme (Fig. 6). With an elevated enzyme concentration, the G3:U70 tRNA^{Cys/CUA} is aminoacylated with alanine, however. This further suggests that the G3:U70 base pair is a major determinant for recognition by the alanine tRNA synthetase.

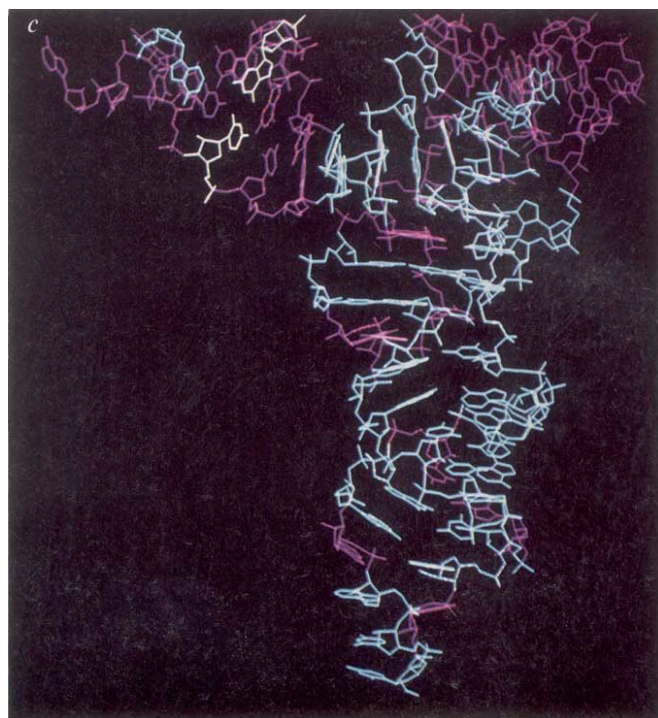
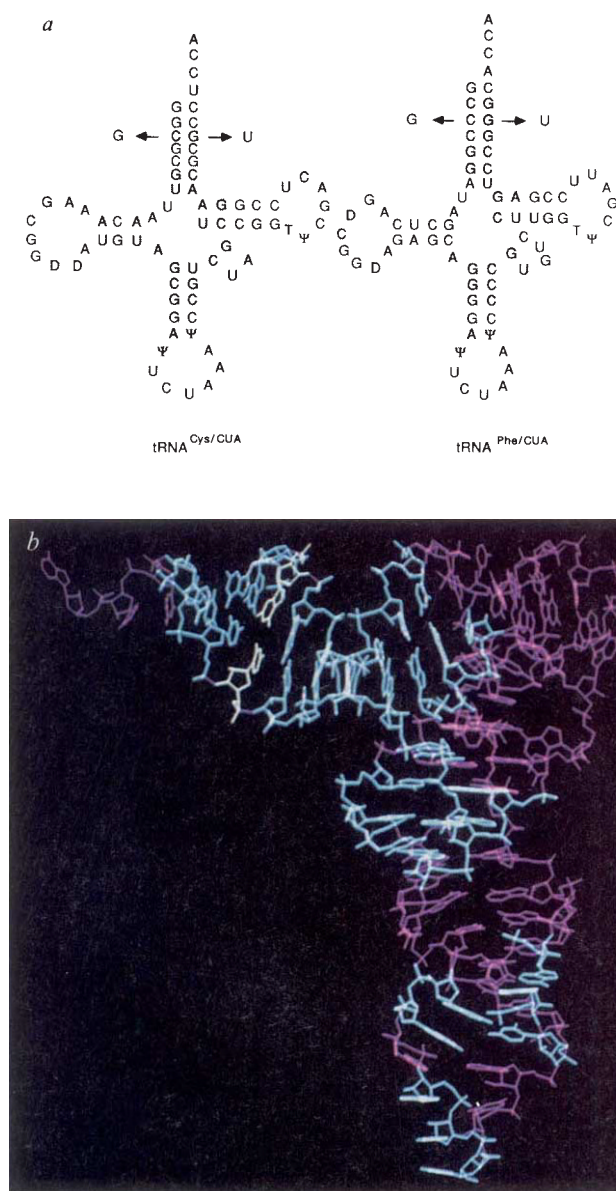


Fig. 5 *a*, Sequences of tRNA^{Cys/ CUA} and tRNA^{Phe/ CUA} (ref. 11). Nucleotides that are different from tRNA^{Ala/ CUA} are shaded. Substitutions at the G3:U70 base pair are indicated by arrows. *b*, Distribution of the 36 positions that have been altered in tRNA^{Ala} (highlighted in blue) in a skeletal model of the three-dimensional tRNA structure. The G3:U70 base pair is highlighted in yellow. For this depiction the sequence of tRNA^{Ala/ CUA} has been built into the known coordinates of the yeast tRNA^{Phe} crystal structure (from the Brookhaven data bank), using the PS300 FRODO programme⁵⁰. *c*, Distribution of nucleotides in tRNA^{Ala} that are different from tRNA^{Cys/ CUA} (highlighted in blue). The G3:U70 base pair is highlighted in yellow. The two nucleotides that are absent in tRNA^{Cys/ CUA} have been treated for this illustration as base differences.

Discussion

Mutations in the acceptor stem of a tRNA^{Tyr} amber suppressor can lead to misacylation with glutamine²⁶⁻³¹. Some of these mutations do not reproduce nucleotides that are in the analogous locations of tRNA^{Gln}. A complicating factor is that the CUA amber-suppressor anticodon can cause some tRNAs, with or without further modification, to be misacylated with glutamine²⁶⁻³³. One possibility is that the glutamine enzyme aminoacylates tRNAs that no other synthetase will aminoacylate⁷. In the present studies, no evidence for insertion of glutamine was found for any of the tRNA amber species in cases where the suppressed *fol* gene product was sequenced (see above).

The mutant alanine tRNAs investigated in this study have base substitutions that are largely concentrated along the inside of the L-shaped tRNA structure (Fig. 5*b*). While substantial alterations to the dihydrouridine, T ψ C and anticodon stems do not prevent aminoacylation with alanine, one of the simplest possible changes—G:U $\rightarrow$ G:C or A:U—in the acceptor helix is sufficient to prevent the tRNA from being a substrate for the alanine enzyme. That this single base pair confers alanine acceptance on tRNA^{Cys} and tRNA^{Phe}, and on the wild-type and mutant tRNA^{Ala/ CUA} sup⁺ species, shows that the G3:U70 base pair is

read by the alanine enzyme in the context of entirely different tRNA sequence frameworks. In the three-dimensional tRNA structure, the distribution of base differences between G3:U70 tRNA^{Cys/ CUA} and tRNA^{Ala/ GGC} (Fig. 5*c*) includes many which are on the opposite side of the structure from those base substitutions which make up the collection of mutant tRNA^{Ala/ CUA} sup⁺ species (Fig. 5*b*). These considerations emphasize the dominant effect of the G3:U70 base pair.

There are probably context effects which affect the efficiency of recognition by the alanine enzyme of a tRNA that has a G3:U70 base pair. We estimate that, under the conditions used, the level of *in vitro* aminoacylation of G3:U70 tRNA^{Cys/ CUA} with alanine is ~20% (Fig. 6). The total level of aminoacylation³⁴ *in vitro* is greatly influenced by the rate of the synthetase-dependent, substrate-independent deacylation (an 'editing' function) reaction³⁵. It is possible that the alanine enzyme has some deacylation activity towards the alanine-aminoacylated G3:U70 tRNA^{Cys/ CUA} species. This deacylation activity may not be as relevant *in vivo*, however, because elongation factor Tu rapidly sequesters aminoacylated tRNA species¹. Exploration of the effect of the G3:U70 base pair in other tRNAs, and further kinetic characterizations of the aminoacylation of the G3:U70 tRNA^{Cys} and G3:U70 tRNA^{Phe} species, will give more insight into context effects.

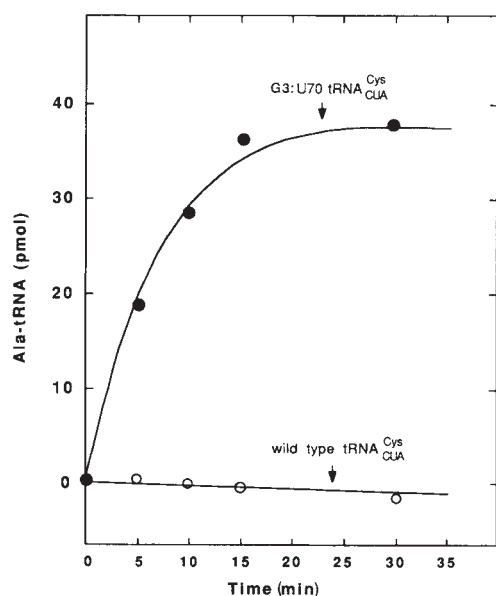


Fig. 6 Aminoacylation activity with alanine of the G3:U70 tRNA^{Cys/Ala} mutant suppressor and of the wild-type counterpart. Assay conditions were similar to those in Fig. 4 and used 4 μ M of each tRNA with 20 μ M of purified alanine-tRNA synthetase. A 10-fold reduction in enzyme concentration also gives aminoacylation of G3:U70 tRNA^{Cys/Ala} (the plateau level is reduced by one-half) with no aminoacylation of tRNA^{Cys/Ala}.

Inspection of published *E. coli* tRNA sequences shows that, while both the tRNA^{Ala/GGC} and tRNA^{Ala/UGC} isoacceptors have a G3:U70 base pair, it is absent from all other *E. coli* tRNAs¹⁹. (An early sequence of tRNA^{Ala/UGC} reported a G3:C70 base pair³⁶ but a subsequent analysis which reported the presence of the G3:U70³⁷ pair has been confirmed by the sequence of the gene³⁸.) This suggests that, in principle, tRNA^{Ala} molecules can be distinguished by an enzyme on the basis of this single base pair in the acceptor helix. This offers encouragement for computer analyses which attempt to determine minimal sequence elements that can be used to distinguish a tRNA from all others in the same organism³⁹. On the other hand, the number of computer-based 'solutions' for a given tRNA is quite large.

As stated above, the determinants of the identity of a tRNA are probably idiosyncratic. In particular, the position 3 to position 70 base pair cannot be the major site of recognition for all of the tRNAs because many tRNAs have the same base pair at that site. The location of major determinants for the identity of

tRNA^{Cys} and tRNA^{Phe} are unknown. That incorporation of the G3:U70 base pair converts the tRNA^{Cys/Ala} suppressor to an alanine-tRNA suggests that recognition of the G3:U70 pair by the alanine enzyme completely outweighs recognition by the cysteine tRNA synthetase. Possibly, positions 3 and 70 are part of the recognition site for the cysteinyl enzyme, and/or the G3:U70 substitution has significantly perturbed the interaction with that enzyme. On the other hand, while the G3:U70 substitution confers ability for alanine acceptance onto tRNA^{Phe/Ala}, there is still some aminoacylation *in vivo* with phenylalanine. It will be necessary to investigate a series of variant tRNA^{Phe} molecules in order to identify a determinant that can be altered so as to eliminate acceptance by phenylalanine. In the context of that hypothetical variant, the G3:U70 base pair might be aminoacylated only by alanine.

A mutant lysine tRNA has been isolated that retains specificity for lysine codons but inserts alanine or glycine⁴⁰. The mutant missense suppressor has a single nucleotide substitution which creates a G3:U70 base pair in the acceptor helix. While it is not established whether this tRNA is misacylated with alanine or glycine, the authors argue for alanine⁴⁰. The results reported here support this conjecture.

The G:U base pair is of the wobble configuration⁴¹. It is relatively common in tRNA secondary structures, although there is usually no more than one per molecule¹⁹. The G:U pairs occur in various parts of the tRNA cloverleaf and could, in principle, be used to distinguish at least some tRNA molecules from each other. In the three-dimensional crystal structure of yeast tRNA^{Phe}, a G4:U69 pair is located in the acceptor stem where it is centred somewhat off the helix axis⁴²⁻⁴⁵. With a G:U wobble pair, the exocyclic 2-amino group of guanosine is not hydrogen bonded and projects into the minor groove, while the free 4-keto group of uracil is in the major groove. The subtle features of the G3:U70 base pair in tRNA^{Ala} are evidently recognized by the alanine enzyme so that G:U is distinguished from other base pairs at this position.

This work was supported by the NIH. We thank John Abelson and Jeffrey Miller for plasmids encoding suppressor tRNAs and the mutant *fol* gene, and for discussions, Emanuel Murgola for *E. coli* strain FTP3689 and for discussion and technical help, Professor U. L. RajBhandary for tRNA-PAGE, Professor Matsudaira for data on protein sequences, Jennifer Normanly for helpful experimental suggestions and Shin-Fuh Lin and Robert Davenport for assistance with the molecular graphics. We are grateful to John Abelson, Helen Edwards, William McClain, Emanuel Murgola, Jennifer Normanly, LaDonne Schulman, Ruth Starzyk and Olke Uhlenbeck for comments on the manuscript.

Received February 16; accepted 14 April 1988.

- Schimmel, P. R. & Soll, D. A. *Rev. Biochem.* **48**, 601-648 (1979).
- Kisselev, L. *Prog. Nucleic Acid Res. molec. Biol.* **32**, 237-266 (1985).
- Schimmel, P. A. *Rev. Biochem.* **56**, 125-158 (1987).
- Knollton, R. G., Soll, L. & Yarus, M. *J. molec. Biol.* **139**, 705-720 (1980).
- Schulman, L. H. & Pelka, H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6755-6759 (1983).
- Samson, J. & Uhlenbeck, O. C. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1033-1037 (1988).
- Normanly, J., Odgen, R. C., Horvath, S. J. & Abelson, J. *Nature* **321**, 213-219 (1986).
- Jasin, M., Regan, L. & Schimmel, P. *Nature* **306**, 441-447 (1983).
- Jasin, M., Regan, L. & Schimmel, P. *Cell* **36**, 1089-1095 (1984).
- Regan, L., Bowie, J. & Schimmel, P. *Science* **235**, 1651-1653 (1987).
- Normanly, J., Masson, J.-M., Kleina, L., Abelson, J. & Miller, J. H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6548-6552 (1986).
- Masson, J. M. & Miller, J. H. *Gene* **47**, 179-183 (1986).
- Nakamura, K. & Inouye, M. *Cell* **18**, 1109-1117 (1979).
- Mims, B. H., Prather, N. E. & Murgola, E. J. *J. Bact.* **162**, 837-839 (1985).
- Rafferty, L. A., Yarus, M. *EMBO J.* **6**, 1499-1506 (1987).
- Rich, A. & Schimmel, P. R. *Nucleic Acids Res.* **4**, 1649-1665 (1977).
- Crothers, D. M., Seno, T. & Soll, D. G. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3063-3067 (1972).
- Murgola, E. J. & Hijazi, K. A. *Molec. gen. Genet.* **191**, 132-137 (1983).
- Sprinzel, M., Hartmann, T., Meissner, F., Moll, H. & Vorderwulbecke, T. *Nucleic Acids Res.* **15**, r53-r188 (1987).
- Murgola, E. J. A. *Rev. Genet.* **19**, 57-80 (1985).
- Mosteller, R. D., Goldstein, R. V. & Nishimoto, K. R. *J. biol. Chem.* **252**, 4527-4532 (1977).
- Yanofsky, C. & Crawford, I. P. *Enzymes* **7**, 1-31 (1972).
- Baccanari, D., Philips, A., Smith, S., Sinski, D. & Burchall, J. *Biochemistry* **14**, 5267-5273 (1975).
- Seong, B. L. & RajBhandary, U. L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 334-338 (1987).

- Jasin, M., Regan, L. & Schimmel, P. *J. biol. Chem.* **260**, 2226-2230 (1985).
- Hooper, M. L., Russell, R. L. & Smith, J. D. *FEBS Lett.* **22**, 149-155 (1972).
- Shimura, Y. *et al. FEBS Lett.* **22**, 144-148 (1972).
- Smith, J. D. & Celis, J. E. *Nature new Biol.* **243**, 66-71 (1973).
- Celis, J. E., Hooper, M. L. & Smith, J. D. *Nature new Biol.* **244**, 261-264 (1973).
- Ghyben, A. & Celis, J. E. *J. molec. Biol.* **83**, 333-351 (1974).
- Inokuchi, H., Celis, J. E. & Smith, J. D. *J. molec. Biol.* **85**, 187-192 (1974).
- Yaniv, M., Folk, W. R., Berg, P. & Soll, L. *J. molec. Biol.* **86**, 245-260 (1974).
- Schulman, L. H. & Pelka, H. *Biochemistry* **24**, 7309-7314 (1985).
- Dietrich, A., Kern, D., Bonnet, J., Giege, R. & Ebel, J.-P. *Eur. J. Biochem.* **70**, 147-158 (1976).
- Schreier, A. A. & Schimmel, P. *Biochemistry* **11**, 1582-1589 (1972).
- Williams, R. J., Nage, W., Roe, B. & Dudock, B. *Biochem. biophys. Res. Commun.* **60**, 1215-1221 (1974).
- Lund, E. & Dahlberg, J. E. *Cell* **11**, 247-262 (1977).
- Young, R. A., Macklis, R. & Steitz, J. A. *J. biol. Chem.* **254**, 3264-3271 (1979).
- McClain, W. H. & Nicholas, H. B. Jr *J. molec. Biol.* **194**, 635-642 (1987).
- Prather, N. E., Murgola, E. J. & Mims, B. H. *J. molec. Biol.* **172**, 177-184 (1984).
- Crick, F. H. C. *J. molec. Biol.* **19**, 548-555 (1966).
- Kim, S.-H. *et al. Science* **185**, 435-440 (1974).
- Robertus, J. D. *et al. Nature* **250**, 546-551 (1974).
- Ladner, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 4414-4418 (1975).
- Quigley, G. J. & Rich, A. *Science* **194**, 796-806 (1976).
- Dente, L., Cesareni, G. & Cortese, R. *Nucleic Acids Res.* **11**, 1645-1655 (1983).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134**, 1141-1156 (1978).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *In Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Bush, B. L., Jones, T. A., Pflugrath, J. W. & Saper, M. A. in *PS300 FRODO-Molecular Graphics Program For The PS300* (ed. Sack, J. S.) (PS300 FRODO Version 6.4, 1987).

On the rotation axis of comet Halley

H. U. Keller & N. Thomas

Max-Planck-Institut für Aeronomie, PO Box 20,
D-3411 Katlenburg-Lindau, FRG

The observations of comet Halley by the Halley Multicolour Camera (HMC) on board the European Space Agency's Giotto spacecraft were supported by an extensive ground-based observational programme using the same filter set¹. The intention was to place the transitory but detailed observations of the encounter in the frame of ground-based observations, thereby yielding the evolution of the comet. The importance of this connection has increased because of the limited information on the three-dimensional extent of the cometary nucleus and its environment retrievable from HMC observations, caused by the communication loss just before closest approach. Here we model dust jets visible on the ground-based images, assuming active sources on the rotating nucleus. We found no fit using the rotation axes so far derived from nucleus observations. A mean instantaneous spin axis, at the time of the Giotto encounter, pointing towards $\alpha = 328^\circ$, $\delta = -63^\circ$ is consistent with both ground-based and HMC observations of the dust. We have confirmed a rotation period of about two days.

Figure 1a, c, and d shows one high-resolution image of comet Halley taken at the South African Astrophysical Observatory (SAAO)² and two images taken at the Anglo-Australian Observatory (AAO)³. All images have been reduced and calibrated, and are part of the HMC database. The scales are 390 and 345 km per pixel, respectively. Azimuthal gradients on the images are emphasized by rotation and subtraction of images (see ref. 4). Dark parts of the images represent maxima within

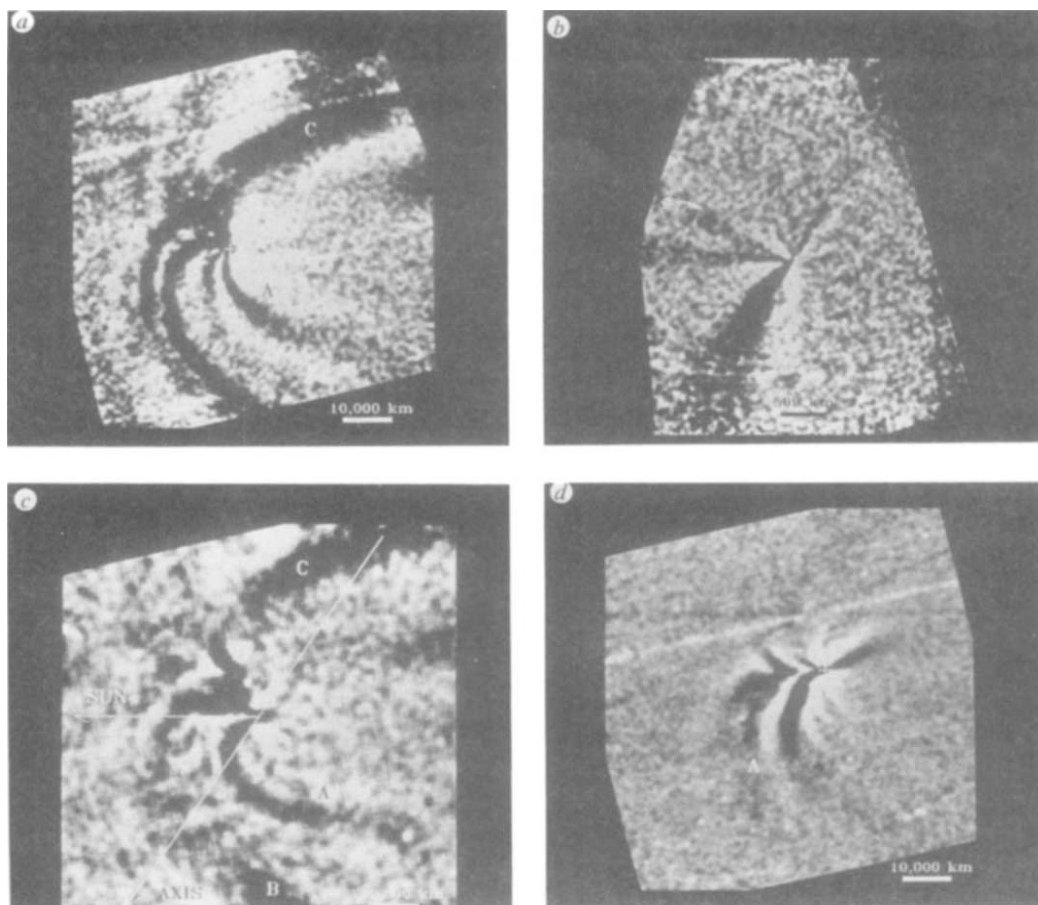
the jet emissions. At least three jets are discernible in Fig. 1c. The southern jet (A) coincides with the strongest emission on HMC images, as conceived by Cosmovici *et al.*¹. The jet can be followed out to $>50,000$ km. It is strongly curved on the scale of the images.

Compare these images with HMC image no. 1586 taken 2 h 8 min before encounter (Fig. 1b), which shows the inner coma of the comet at a resolution of 11.8 km per pixel out to a distance of $\sim 2,300$ km from the nucleus in projection. The strongest emission points about 60° southward of the comet-Sun direction⁵. Detailed investigations have shown that this jet is slightly curved and variable with time, reflecting the rotation of the nucleus⁶. The source strength of this jet does not vary within the accuracy of a few percent in HMC images covering the 3 h before closest approach.

Our aim here is to connect the emissions visible on HMC images with the jets on ground-based images. We assume that the jets are emitted from discrete sources on the surface of the nucleus in a direction characterized by the latitude of the source and a longitudinal angle relative to the assumed rotation axis. The position (the direction) of the source in inertial space changes with the rotation of the nucleus. Initially we assumed the values derived by Wilhelm *et al.*⁷ for the orientation of the rotation axis. From observations of the nucleus by three spacecraft, it was concluded that the axis points towards $\alpha = 50^\circ$ and $\delta = -40^\circ$ in celestial coordinates (1950). The corresponding rotation period is 54 h. We assume the production of the dust source to be constant.

The projected outflow speed of the dust in the jet can be measured directly by determining the motion of the jet perpendicular to the comet-Sun line on observations taken at different times. On South African images taken 2 h apart, the line of the jet (A) south of the Sun-comet line (Fig. 1c) moves away from the nucleus with a projected outflow velocity of 400 ± 50 m s⁻¹,

Fig. 1 Observations of comet Halley over a 24-h period around the time of Giotto encounter. Each image has been treated using an azimuthal gradient technique to reveal local maxima, which appear dark. In all cases, the projection of the Sun-comet line is precisely to the left and ecliptic north is up. The projection of the rotation axis derived by Wilhelm *et al.*⁷ is indicated in c. a, Image taken from Australia 6 h before Giotto closest approach; b, HMC image 1586 taken 2 h before closest approach; c, image taken from South Africa 1.5 h after closest approach; d, image from Australia, 18 h after closest approach.



yielding a lower limit for the actual outflow velocity. By extrapolation, the position of the jet on the Australian images taken 8 h earlier and 16 h later can be predicted, allowing positive identification of the same feature on all images.

Figure 2 displays the geometrical relationship between the lines of sight of HMC and from Earth at the time of the Giotto encounter. The corresponding phase angles (Sun-comet-observer) are 107.2° and 63.6° . The angle between the relative velocity vector (along which HMC looked) and the comet-Earth line is 44.3° . This vector is inclined by 10° north relative to the ecliptic. The rotation axis of Wilhelm *et al.*⁷ is fortuitously in the plane of observation when viewed from Earth (Fig. 2).

Figure 3 gives a schematic representation of a jet and a rotation axis in the image plane. The sense of rotation is the same as for the ground-based observations of comet Halley (that is, direct with respect to the orbit). The axis is inclined at 34° to the vertical. The angle, α , can only be positive (for emission at constant velocity) and depends on the latitude of ejection and the ratio of the rotational to the emission velocity. In addition, when the projection of the jet crosses the axis, there is a change in the sense of curvature of the jet. Comparison with the observation in Fig. 1c shows that the curvature of jet A does not change: it retains its sense far beyond the projected rotation axis. This is a clear indication of an incompatibility with the assumed rotation axis. The above inference has been confirmed by repeated but unsuccessful attempts to model the curvature of the jet with this axis.

We therefore abandoned the suggested rotation axis. An excellent fit for jet A has been achieved with the rotation axis pointing towards $\alpha = 328^\circ$ and $\delta = -63^\circ$. The axis determined by comparing HMC and Vega images is 52° from this. The outflow speed, v_a , has been found to be 420 m s^{-1} , in good agreement with the measured projected speed. The curvature of the jet is hardly influenced by the effects of solar radiation pressure; a value of 0.35 ± 0.15 for β , the ratio of the force from solar radiation pressure to that from solar gravity, was found to be in agreement with the data: we have used $\beta = 0.35$. The influence of the solar radiation pressure force is therefore of only minor importance within 20,000 km of the nucleus. The source has a latitude of 32° north on the nucleus (in southern ecliptical direction). Its longitude is such that it points 49° out of the HMC image plane, away from the observer at the time of closest approach (and 48° south of the comet-Sun line in projection). The source passed the morning terminator during the morning of 13 March 1986. The farthest observed parts of the jet are in agreement with this result. We therefore predict that jet A will not be visible on pre-encounter HMC images at 09:00 UT on 13 March 1986.

The course of jet A as observed on the ground-based images is marked as a dotted line on the diagrams of Fig. 4 corresponding to the images in Fig. 1. The fit is remarkably good and well within the resolution of the images. The transformation into the observational plane of HMC yields a projected direction of 55° south of the comet-Sun line at the time of the first transmitted full-frame image 3 h before closest approach (see Fig. 4b) and a curvature similar to that observed.

We have attempted preliminary fits to other jets visible on the ground-based images. For the strongly curved jet towards the Sun (B), a reasonable fit could be found using $\beta = 0.35$ and $v_a = 400 \text{ m s}^{-1}$. The jet had already turned off before the start of night operations of HMC. The linear appearance of the straight jet pointing almost anti-Sunward (Fig. 1d) is due to projection effects and is not caused by emission from a constantly illuminated pole.

The rotation axis determined from comparison of nucleus images achieved by the Vega and Giotto spacecraft does not fit the evolution of dust jets during the time span of the Giotto encounter. Other determinations of the orientation of the axis from the spacecraft observations have yielded similar values to that found by Wilhelm *et al.*⁷ (see ref. 8 for a recent summary and discussion). The obliquities of the rotation axes for all these

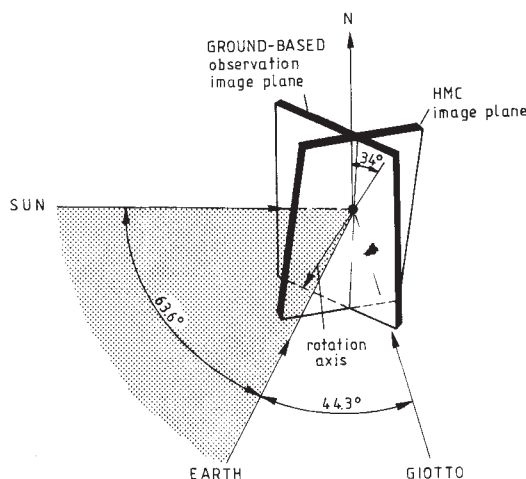


Fig. 2 Observing geometry for the ground-based and Giotto observations. The rotation axis derived by Wilhelm *et al.*⁷ is indicated.

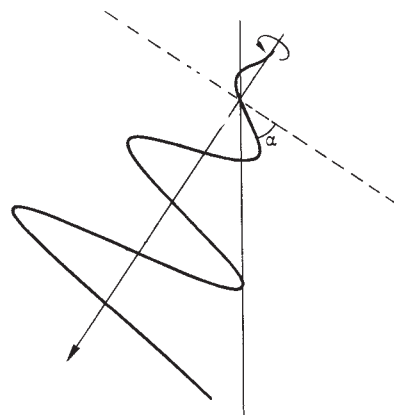


Fig. 3 Schematic diagram of the evolution of a jet emitted from a rotating source, the rotation axis of which is in the image plane.

orientations are $\sim 30^\circ$. This is also true of the axis suggested here: its obliquity is 28° , although its orientation differs from the old value by more than 50° . In earlier images, taken 2 days before encounter, better fits are obtained using a rotation axis differing by about 20° in orientation. This could indicate the presence of nutation (free precession) as widely assumed to reconcile the short and long periods^{9,10}.

The dynamics of the jets and their curvatures do not conform to a long rotation period of about 7.4 days as suggested from the periodicity of the gas and dust production rates^{8,11}. The observed outflow speeds require rotation periods of ~ 2 days. The exact value has not been determined here because the period has not been treated as a free parameter.

A recalculation and confirmation of the rotation axis and period from the spacecraft observations is required. It must be checked whether the present axis orientation could be brought into accordance with the orientations of the nucleus during the two Vega and the Giotto encounters, perhaps using a slightly modified rotation period and nutation. At present there is a clear discrepancy between the derivation from the dust jets and from the spacecraft observations.

The observational programme of the HMC experiment on the spacecraft and on the ground has yielded unprecedented observations of dust jets from the surface of the nucleus out to distances of $> 50,000 \text{ km}$. The results presented in this paper demonstrate that important parameters, not only of the dust and its dynamics but also of the orientation and dynamics of

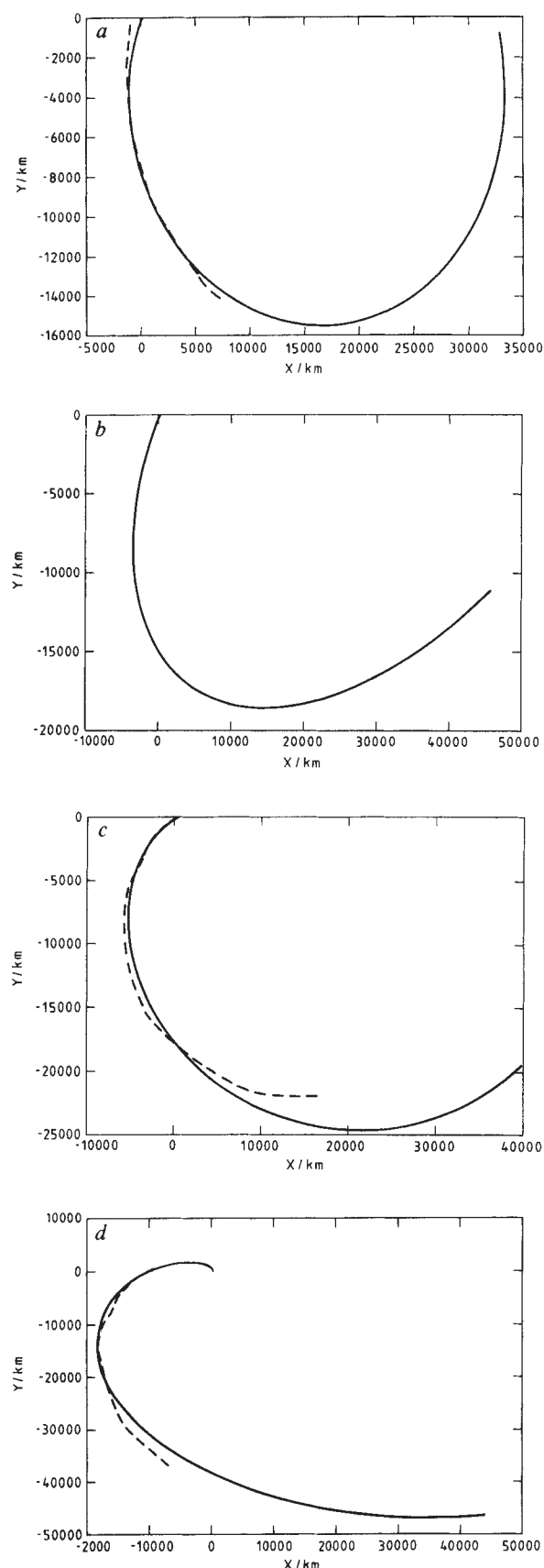


Fig. 4 Model calculations of the evolution of the south jet. Frames *a*–*d* correspond to Fig. 1*a*–*d*. The dotted line indicates the line of the local maximum; the solid line gives the model result using the values given in the text. In *b* the transformation into the HMC system is given. No dotted line is given as the HMC observation was at a very much smaller scale.

the nucleus itself, can be derived from careful study of the entire data set.

We thank our colleagues of the HMC team for many discussions and encouragement, and particularly the observers whose data was made available so quickly and was used for the interpretation in this paper: Drs C. B. Cosmovici and P. Mack for the SAAO and Drs D. W. Hughes and S. F. Green for the AAO images and their calibration.

Received 13 January; accepted 25 February 1988.

1. Cosmovici, C. B. *et al. Eur. Space Ag. spec. Publ.* **250** (II), 375–379 (1986).
2. Cosmovici, C. B., Mack, P., Craubner, H. & Schwarz, G. *Eur. Space Ag. spec. Publ.* **250** (II), 151–155 (1986).
3. Green, S. F. & Hughes, D. W. *Eur. Space Ag. spec. Publ.* **250** (II), 157–161 (1986).
4. Thomas, N. & Keller, H. U. *Astr. Astrophys.* **187**, 843–846 (1987).
5. Keller, H. U. *et al. Astr. Astrophys.* **187**, 807–823 (1987).
6. Thomas, N. & Keller, H. U. *Bull. Am. astr. Soc.* **19**, 866 (1987).
7. Wilhelm, K. *et al. Eur. Space Ag. spec. Publ.* **250** (II), 367–369 (1986).
8. Sekanina, Z. *Comet Halley 1986: World-wide investigations, Results and Interpretations* (Ellis Horwood, Chichester) (in the press).
9. Sekanina, Z. *Nature* **325**, 326–328 (1987).
10. Julien, W. M. *Nature* **326**, 57–58 (1987).
11. Millis, R. L. & Schleider, D. G. *Nature* **324**, 646–649 (1986).

Surficial textures of the Galilean satellites

Bonnie J. Buratti, Robert M. Nelson & Arthur L. Lane

Jet Propulsion Laboratory, California Institute of Technology,
4,800 Oak Grove Drive, Pasadena, California 91109, USA

Knowledge of the textural characteristics of planetary surfaces is one of the objectives of remote sensing observations. The comparison of accurate photometric measurements with scattering models yields estimates of the compaction state of the optically active portion of the regoliths of airless bodies. We have analysed as a function of solar phase angle the observations of the Galilean satellites of Jupiter obtained with the International Ultraviolet Explorer^{1,2}. By fitting the measurements to a shadowing model³, comparative descriptions of the microtextures of the optically active portion of the surfaces of the satellites are derived. Important differences among the satellites, and between leading and trailing hemispheres of individual satellites, result from the different processes of meteoritic bombardment, magnetospheric interaction and geological resurfacing that operate in the Jovian system. Io and Callisto have the most tenuous upper regoliths, whereas the surface of the leading side of Europa is the most compact.

Ultraviolet (UV) observations of surfaces are particularly useful, because in this region of the spectrum multiple scattering (one of the most difficult factors to incorporate into a scattering model) is negligible. The comparison of laboratory measurements of the scattering properties of geological samples of known compaction states with remote sensing observations and theoretical models provides a third, and more concrete, basis on which to derive the compaction state of the surface. Comparison of the textures of the surfaces of various planets and satellites places important constraints on the different evolutionary processes at work on their surfaces.

We report on an analysis of the first acquisition of photometrically accurate UV measurements of the Galilean satellites from the International Ultraviolet Explorer (IUE) satellite^{1,2}. The observations were obtained between 1978 and 1985. We also compare our results with laboratory directional scattering properties of geological samples of both high and low compaction states.

The IUE data are integrated over the spectral range 0.24–0.27 μm (centre at 0.26 μm). For each of the four Galilean

satellites, the integrated brightness is further binned into the geographical domains of 'leading side' (corresponding to geographical longitudes 45°–135°) and 'trailing side' (corresponding to 225°–315°). Finally, the observations are represented as a function of solar phase angle (the angle between the Earth, Sun and satellite) to create a solar phase curve of both sides of the satellites. These curves express how the satellite scatters radiation as a function of viewing geometry. From Earth, the excursion in phase angles is limited by geometric considerations to the range 0–12° (Voyager observations were obtained at higher phase angles; however the UV filter on the imaging experiment was centred at 0.35 μm). Variations in the IUE observations as a function of geographical longitude are given in ref. 2.

Recent radiative transfer models of the surfaces of airless bodies^{4–8} have described the bidirectional scattering properties of these surfaces in terms of four physical properties: the single scattering albedo, the compaction state of the optically active portion of the upper regolith, the single-particle phase function, and the degree of macroscopic roughness. The single scattering albedo controls the degree of multiple scattering on the surface. For the low UV geometrical albedos of the Galilean satellites (Table 1), both laboratory measurements and theoretical considerations demonstrate that the effects of multiple scattering are negligible^{9,10}. Similarly, the effects of macroscopic roughness (due to features ranging in size from clumps of several particles to mountains and craters) can be ignored between the phase angles of 0° and 12°; this effect becomes important for phase angles beyond ~30° (refs 5, 11). For the range in phase angles over which the IUE data were obtained, the integrated brightness depends primarily on the compaction state of particles comprising the optically active portion of the upper regolith^{3,5,12}.

It is well known that most airless planets and satellites exhibit an anomalous increase in brightness as their surfaces become fully illuminated to the observer (reviewed in ref. 13 for the Galilean satellites). This so-called opposition effect becomes most prominent at solar phase angles <6°. The conventional explanation of the effect is that mutual shadowing among the particles in the regolith disappears as the phase angle approaches 0°. An understanding of the microstructure of the upper regolith of a planetary surface can be gained by fitting observations to a theoretical model expressing the reflected radiation as a function of solar phase angle. Analysis of ground-based visible observations has always been a problem because of the difficulty of fully accounting for the role of multiple scattering.

We have fitted the shadowing model of Irvine³ to each set of observations of the Galilean satellites by a non-linear least-squares method. The model of Irvine assumes that the constituent particles on the surface are single-sized and spherical. Computations by Kawata and Irvine¹⁴ showed that including the effects of a non-uniform size distribution does not affect the results of Irvine's model, as long as the distribution does not vary sharply and the particles are all within the geometrical optics limit. The shape of the solar phase curve and the ampli-

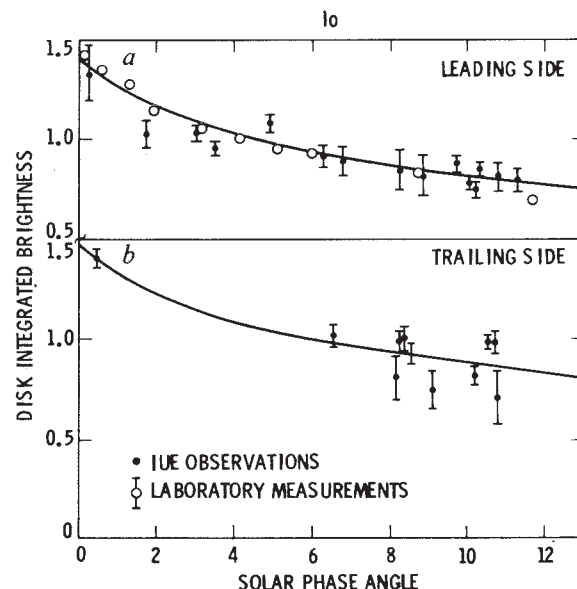


Fig. 1 IUE observations of the *a*, leading and *b*, trailing sides of Io. The lines are the best-fit theoretical representations of Irvine's shadowing model, corresponding to a model fractional void space of 87%. Also included are laboratory measurements of a fluffy sample of particulate basalt (open circles) with a measured void space of 90%. All data are normalized to unity at 6°.

tude of the opposition effect are determined by a model compaction parameter defined as

$$D = \frac{3}{4\pi} \frac{\rho}{\rho_0}$$

where ρ is the bulk density of the optically active portion of the regolith, and ρ_0 is the density of the individual particles. For porous, fluffy surfaces where shadowing is more significant, the magnitude of the opposition effect is higher.

The results of our investigation are listed in Table 1 and discussed for each satellite. We have found that for a Henyey-Greenstein single-particle phase function, which is a commonly used empirical function for planetary surfaces¹², our results are insensitive to the selection of the asymmetry factor describing the forward or backward scattering component of the reflected radiation. Thus our measurements, because they were obtained at small solar phase angles and at a wavelength for which the geometrical albedos of the Galilean satellites are low, are uniquely sensitive to the compaction state of the upper regolith of the bodies.

Io. We have previously reported that the leading side of Io is extremely porous, on the basis of a preliminary analysis¹. Our best fit to the data is $D = 0.03$, which corresponds to a model void space (the fraction of the bulk volume not occupied by particles) of about 87%. Although the observations for the trailing side (see Fig. 1) have large errors bars, the measurements are consistent with a similarly high porosity. We thus believe that the entire surface of Io is very fluffy, consistent with a texture that is the result of the deposition of volcanic particles.

We have measured the scattered intensity of a laboratory sample of particulate fluffy basalt as a function of viewing geometry. The sample of basalt has a measured void space of 90% and a normal reflectance of 0.24 (this value is higher than that of Io, but still in the region where multiple scattering is not important). Figure 1 shows that the laboratory measurements are in good agreement with our IUE observations and the theoretical model.

Our results for Io agree with an analysis by Simonelli and Veverka¹⁵; they fit Voyager observations of Io to Hapke's recent improved model of the opposition effect⁶, and they obtained a

Table 1 Differences among satellites and between their leading and trailing hemisphere

Satellite	Geometric albedo ² (0.26 μm)	Compaction parameter (D)	Model % void space ($1 - \rho/\rho_0$)
J1 Io (L)	0.015 ± 0.001	0.03 ± 0.01	$87 \pm 5\%$
(T)	0.028 ± 0.002	0.03 ± 0.02	$87 \pm 8\%$
J2 Europa (L)	0.180 ± 0.004	$0.18^{+0.01}_{-0.04}$	$25^{+17\%}_{-5\%}$
(T)	0.096 ± 0.002	0.05 ± 0.01	$79 \pm 4\%$
J3 Ganymede (L)	0.128 ± 0.007	0.04 ± 0.01	$83 \pm 4\%$
(T)	0.070 ± 0.003		
J4 Callisto (L)	0.040 ± 0.008	0.02 ± 0.01	$92 \pm 5\%$
(T)	0.056 ± 0.002		

L, Leading side; T, Trailing side.

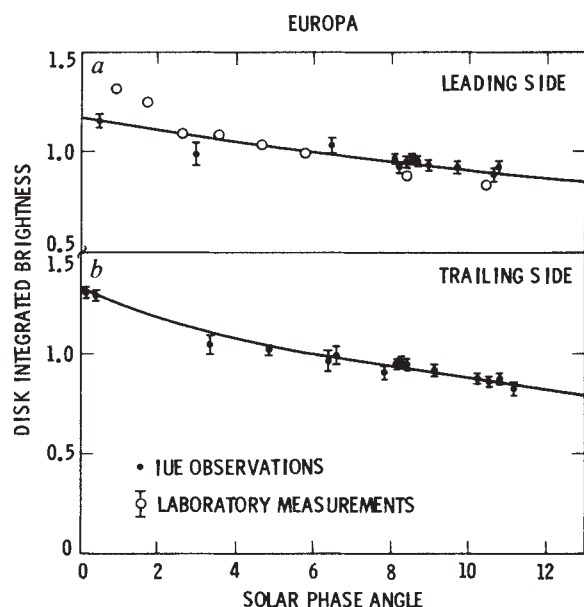


Fig. 2 IUE observations of the *a*, leading and *b*, trailing sides of Europa. The best-fit lines correspond to 25% (leading) and 79% (trailing) model void space. *a*, Laboratory measurements of particulate basalt, compacted to 25% void space, are also shown (open circles). Although there is agreement between the laboratory measurements and the other two sets of data for phase angles greater than 3°, the laboratory data show a definite opposition surge.

similarly large porosity. We find it hard to reconcile our results with those of Pang *et al.*¹⁶, who report a porosity of 63%, based on a fit of both ground-based and Voyager photopolarimeter data to Lumme's model.

Other lines of investigation have supported an extremely porous surface for Io. Eclipse observations showed that Io loses heat rapidly, which would be consistent with a surface with large void spaces¹⁷. Matson and Nash¹⁸ have presented a cold trapping model for Io's surface which requires about 90% void space.

Europa. The observations for Europa are quite good for both the leading and trailing side (Fig. 2). For the leading side, we have found that the surface is extremely compacted, with only ~25% void space. This side of Europa is the only example we could find of a planetary surface with no opposition effect. Millis and Thompson¹⁹ were the first observers to note this unusual aspect of Europa, and Voyager observations²⁰ detected no opposition effect throughout the 0.35–0.59 μm region, down to a solar phase angle of 3°. Figure 2 also shows our laboratory measurements of a compacted particulate basalt sample with 25% void space. There is good agreement between the three sets of data, except that the laboratory data show a definite opposition surge. This discrepancy could possibly be due to the fact that the surface of Europa does not consist of particulate material.

The IUE observations for the trailing side of Europa show the first evidence of a small opposition surge for that satellite. The fit of the observations to the model corresponds to a surface with 79% void space.

These results are in reasonable agreement with an analysis of Voyager images of Europa in terms of Hapke's function¹¹; the average void space was calculated to be 65% (this result for Europa was significantly lower than that of the Moon and three icy Saturnian satellites).

This analysis shows that there are markedly different surficial microtextures on the leading and trailing sides of Europa. As early as 1975, Millis and Thompson¹⁹ noted that the trailing side has a larger phase coefficient than the leading side, but this difference in ground-based visual phase curves could have been

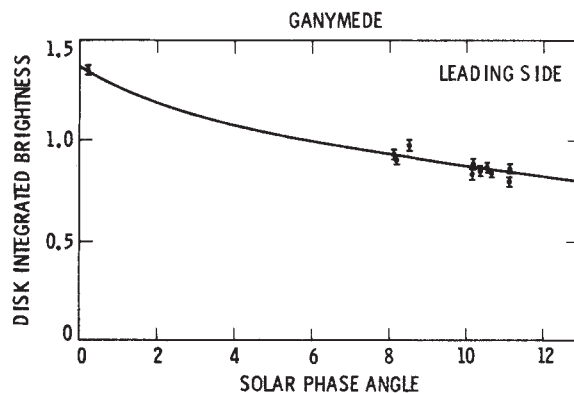


Fig. 3 IUE observations of the leading side of Ganymede with the shadowing model corresponding to 83% void space (observations on the trailing side did not extend over a sufficient range in phase angles to fit a model reliably).

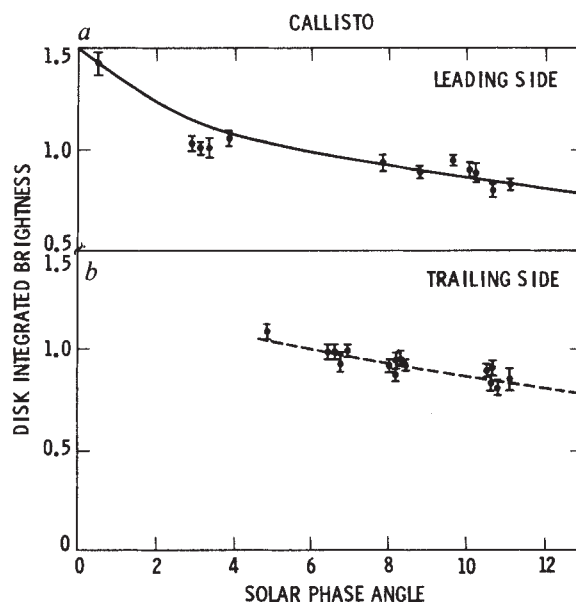


Fig. 4 IUE observations of both hemispheres of Callisto. *a*, Leading side, corresponding to 92% model void space; *b*, the trailing side; observations cannot be reliably fitted to a model.

explained by the diminished importance of multiple scattering on the lower albedo trailing side of Europa. For the IUE observations, multiple scattering is not important for either side; we can thus attribute the difference in the phase curves to structural differences. Clark *et al.*^{21,22} suggested that micrometeoritic bombardment and sputtering could produce leading/trailing asymmetries in the grain sizes of particles, which in turn would be photometrically detectable.

The three possible reasons for the difference in textural properties between the two sides of Europa are: (1) preferential bombardment of the trailing side by micrometeorites²³; (2) interaction of the trailing side of the satellite with slow ions and subsequent implantation of ions and sputtering^{21,24–26}; and (3) unequal distribution of geological units on the two hemispheres²⁷. A leading/trailing asymmetry in colour has been attributed to the second reason²⁸; all three reasons could be contributing factors in producing textural differences. Both sides of Europa are geologically recent²⁷; in the absence of significant amounts of micrometeoritic gardening or magnetospheric sputtering, the leading side would have a texture similar to that of a recent ice flow. For Io we do not see similar textural differences

between hemispheres because the deposition of volcanic particulates occurs on a much more rapid time scale.

Ganymede. The IUE observations of Ganymede contain measurements below 3° only for the leading side of the satellite. This side of Ganymede has a sizeable opposition effect which is smaller than that of Io, but larger than that of both sides of Europa, and corresponds to a void space of 83%. The surface of Ganymede is divided between pristine cratered terrain, and resurfaced grooved terrain²⁹. Because the grooved terrain has been exposed to meteoritic gardening for a shorter time, we would expect its texture to be less fluffy. Thus the compaction state of the upper surface of Ganymede (at least for the leading side) is intermediate between the compact, geologically recent surface of Europa, and the very fluffy, unresurfaced Callisto (see below).

Callisto. IUE observations below 3° exist only for the leading side of Callisto, although coverage on the trailing side between 11° and 3° is good (Fig. 4). The leading side shows a remarkable opposition effect, the largest of the four Galilean satellites. Our observations are in agreement with visible ground-based work^{13,19}. The telescopic measurements detected a larger opposition effect on the leading side: we do not have the observations necessary on the trailing side below 3° to detect hemispherical differences. We find a surface model void space of about 92%. Because Callisto has never undergone any resurfacing and has no large asymmetries in the distribution of geological units, this tenuous structure is probably due to micrometeoritic bombardment (significant effects from ion bombardment are unexpected at Callisto's distance from Jupiter).

In summary, the Galilean satellites have important differences in the textural properties of the optically active portion of their surfaces. Io and Callisto have porous, fluffy surfaces due (in the former case) to the deposition of volcanic particles, and (in the latter case) to eons of micrometeoritic bombardment on a primordial surface. Europa, with recent geological resurfacing, has a compact surface. The trailing side is less compacted due to the gardening effects of micrometeorites and bombardment by slow ions. The surface of Ganymede, with both primordial and resurfaced geological units, is in an intermediate state of compaction. We note that the shadowing model is derived under certain necessary simplifying assumptions. For this reason, it cannot be claimed that our estimates of the compaction state of the surface represent exact physical values. Rather, we aim to make comparisons of this derived model-dependent parameter among the four Galilean satellites to understand differences in their surfaces.

We note that our observations between phase angles of 2° and 4° are consistently below the line for the shadowing model (see Figs 1–4). We do not see a similar pattern in visible phase curves¹³. This discrepancy suggests that the opposition surge may be more sharply peaked in the UV. For the UV range of the spectrum, the size of particle for which the laws of geometrical optics apply is smaller; this means that a larger number of small particles may be contributing to the back-scattering peak in the UV. Either Irvine's shadowing model breaks down because the distribution in particle sizes is too wide, or because physical phenomena in addition to geometrical shadowing are contributing to the shape of the phase curve. The latter might include Mie scattering or glories. Clearly, more observations of the Galilean satellites and other airless bodies are required to determine whether there is a wavelength dependence to the shape or amplitude of the opposition effect.

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology under contract with NASA. We thank B. Hapke and J. Veverka for detailed reviews.

Received 5 October 1987; accepted 25 January 1988.

1. Nelson, R. M. *et al.* *Bull. Am. astr. Soc.* **16**, 684–685 (1984).
2. Nelson, R. M. *et al.* *Icarus* **72**, 358–380 (1987).
3. Irvine, W. M. *J. geophys. Res.* **71**, 2931–2937 (1966).
4. Hapke, B. *J. Geophys. Res.* **86**, 3039–3054 (1981).
5. Hapke, B. *Icarus* **59**, 41–59 (1984).
6. Hapke, B. *Icarus* **67**, 264–280 (1986).
7. Goguen, J. D. thesis, Cornell Univ. (1981).
8. Lumme, K. & Bowell, E. *Astr. J.* **86**, 1694–1704 (1981).
9. Veverka, J., Goguen, J., Young, S. & Elliot, J. *Icarus* **34**, 406–414 (1978).
10. Buratti, B. *Icarus* **59**, 392–405 (1984).
11. Buratti, B. *Icarus* **61**, 208–217 (1985).
12. Veverka, J. in *Planetary Satellites* (ed. Burns, J.) 171–209 (University of Arizona Press, Tucson, 1977).
13. Morrison, D. & Morrison, N. in *Planetary Satellites* (ed. Burns, J.) 363–378 (University of Arizona Press, Tucson, 1977).
14. Kawata, Y. & Irvine, W. in *Exploration of the Planetary System* IAU Symp. 65 (eds Woszczyk, A. & Iwaniszewska, C.) 441–464 (Reidel, Dordrecht, 1974).
15. Simonelli, D. & Veverka, J. *Icarus* **68**, 503–521 (1986).
16. Pang, K. D., Lumme, D. & Bowell, E. *Proc. lunar planet. Sci. Cont.* **12B**, 1543–1553 (1981).
17. Hansen, O. L. *Icarus* **18**, 237–246 (1973).
18. Matson, D. L. & Nash, D. B. *J. geophys. Res.* **88**, 4771–4783 (1983).
19. Millis, R. T. & Thompson, D. T. *Icarus* **36**, 408–419 (1975).
20. Buratti, B. & Veverka, J. *Icarus* **55**, 93–110 (1983).
21. Clark, R. N., Fanale, F. P. & Zent, A. *Icarus* **56**, 233–245 (1983).
22. Clark, R. N. & Lucey, P. B. *J. geophys. Res.* **89**, 6341–6348 (1984).
23. Shoemaker, E. M. & Wolfe, R. F. in *Satellites of Jupiter* (ed. Morrison, D.) 277–339 (University of Arizona Press, Tucson, 1982).
24. Morrison, D. & Burns, J. A. in *Jupiter* (ed. Gehrels, T.) 991–1034 (University of Arizona Press, Tucson, 1976).
25. Lane, A. L., Nelson, R. M. & Matson, D. L. *Nature* **292**, 38–39 (1981).
26. Johnson, T. V. *et al.* *J. geophys. Res.* **88**, 5789–5805 (1983).
27. Lucchitta, B. K. & Soderblom, L. A. in *Satellites of Jupiter* (ed. Morrison, D.) 521–555 (University of Arizona Press, Tucson, 1982).
28. Nelson, M. L. *et al.* *Icarus* **65**, 129–151 (1986).
29. Shoemaker, E. M., Lucchitta, B. K., Plescia, J. B., Squyres, S. W. & Wilhelms, D. E. in *Satellites of Jupiter* (ed. Morrison, D.) 435–520 (University of Arizona Press, Tucson, 1982).

Magnetic anisotropy in the electron momentum density of iron

Malcolm J. Cooper, Stephen P. Collins,
David N. Timms, Adelbaki Brahmia, Prabhakar P. Kane,
Ronald S. Holt* & David Laundy†

Department of Physics, University of Warwick,
Coventry CV4 7AL, UK

* Neutron Division, Rutherford Appleton Laboratory, Chilton,
Oxfordshire OX11 0QX, UK

† Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK

The existence of magnetic terms in the X-ray scattering cross-section is well established^{1–3}, and the suggestion that X-ray studies of magnetization density would be viable dates from 1970⁴. But despite pioneering diffraction studies with conventional X-ray sources^{5–7}, measurements proved difficult until the advent of synchrotrons producing high X-ray flux. This is because of the weakness of these contributions and the general need to use polarized beams or polarization analysis⁸. In recent years, diffraction studies of rare-earth antiferromagnets⁹ and interfacial magnetic layers¹⁰ undertaken with synchrotron radiation have established the viability of the approach and the magnetic terms in the cross-section. In ferromagnets the demonstration of these effects has been predominantly through inelastic Compton scattering, rather than elastic Bragg diffraction studies. The Compton experiment is interpreted in terms of the electron's momentum distribution¹¹, rather than its more familiar spatial counterpart studied in polarized neutron diffraction experiments. Here we report the measurement of the [111], [110] and [100] directional Compton profiles of the unpaired spin electrons in a single crystal of iron, obtained with a 60-keV circularly polarized beam extracted from the Daresbury storage ring. The data provide a more critical test of band theory than is possible with polycrystalline samples (for which the scattering geometry is much simpler), and indicate the potential of this new technique.

A first-order 'magnetic' contribution to the inelastic scattering arises only if circularly polarized radiation is used. It yields a term which, when compared with the leading charge scattering term, is of magnitude¹ (P_c/mc)S. ($\mathbf{k}_i \cos \phi + \mathbf{k}_s$) where P_c is the degree of circular polarization, S the net electronic spin of the

1. Nelson, R. M. *et al.* *Bull. Am. astr. Soc.* **16**, 684–685 (1984).
2. Nelson, R. M. *et al.* *Icarus* **72**, 358–380 (1987).
3. Irvine, W. M. *J. geophys. Res.* **71**, 2931–2937 (1966).
4. Hapke, B. *J. Geophys. Res.* **86**, 3039–3054 (1981).

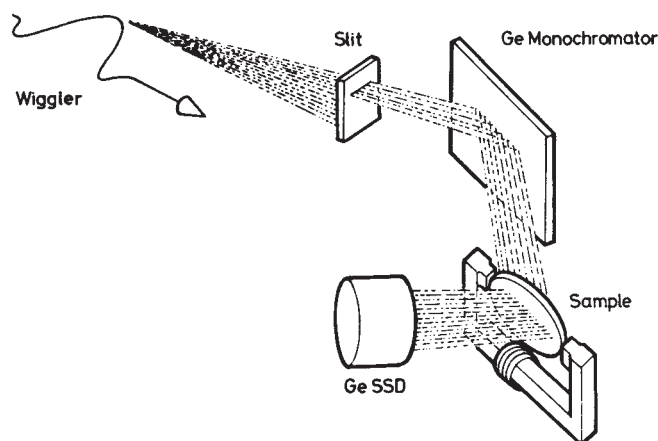


Fig. 1 The Compton scattering experiment. The radiation from the 5 T bend of the wiggler magnet at the Daresbury storage ring emitted at a mean angle of 0.2 mrad to the orbital plane is selected by a slit and monochromated by a plane Ge 220 crystal. The magnetization of the Fe single crystal disc is reversed frequently and the scattered beam detected by the germanium solid-state detector is stored in 'up' and 'down' memories.

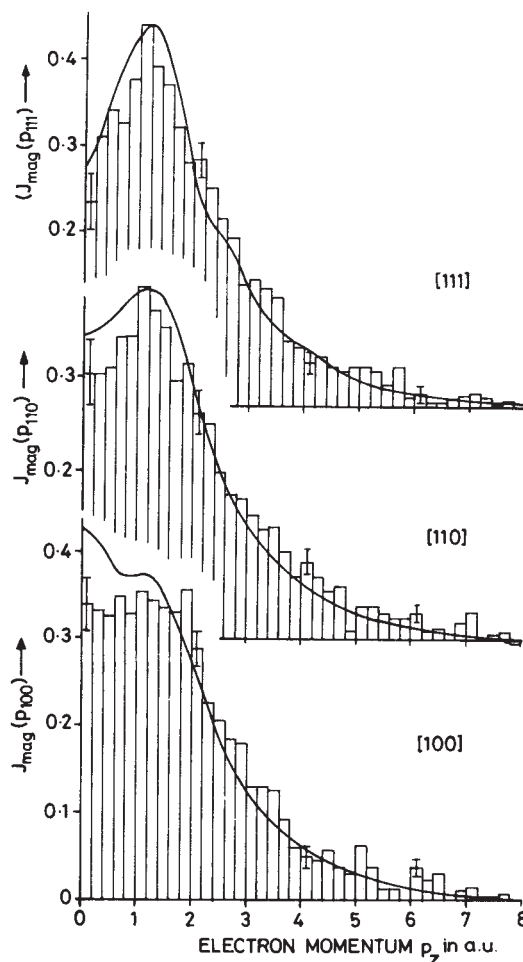


Fig. 2 The histograms show the directional magnetic profiles as a function of electron momentum; the solid line represents the APW calculation convoluted with the experimental resolution function which had a full width at half maximum of 0.7 a.u. ($1.4 \times 10^{-24} \text{ kg m s}^{-1}$). The data collection time for each profile was ~ 20 h and the incident flux of monochromated synchrotron radiation was $\sim 10^8$ photons per second. Only positive p_z values are plotted, the curves being symmetric about $p_z = 0$ and having an area of 2.2 electrons. The magnetic anisotropy is clear; the central dip, which is greatest in the [111] profile, is attributed to a negatively magnetized contribution.

target, $\mathbf{k}_i$ and $\mathbf{k}_s$ the incident and scattered beam wave-vectors, ϕ the angle of scattering, m the electron mass and c the velocity of light. This term scales approximately as the ratio of the photon energy to the rest-mass energy of the electron (511 keV); thus, bearing in mind that more than 90% of the electrons make no net contribution to the magnetic scattering because they have paired spins, the magnetic contribution is never more than a few per cent at photon energies characteristic of the present generation of synchrotrons.

This problem is compounded by the difficulty of obtaining adequate circularly polarized flux at energies that make Compton scattering measurements in transition metals viable, namely 30 keV or above¹². Others have used the circularly polarized gamma radiation emitted by cooled oriented nuclei¹³, where the flux is severely limited by self-heating effects, or they have fabricated X-ray quarter-wave plates to convert from linearly to circularly polarized synchrotron radiation¹⁴. The Warwick group, on the other hand, have developed the simplistic approach of viewing the elliptically polarized radiation which is emitted at a slight angle to the orbital plane of the synchrotron^{12,15}, a method that was first demonstrated in the X-ray regime in a diffraction study of zinc ferrite¹⁶ and also used to measure spin-dependent photoabsorption¹⁷. The use of a high-field superconducting magnet has ensured that a beam with $P_c > \frac{1}{2}$ can be extracted at photon energies of $\sim mc^2/10$ and at a viewing angle $\sim \frac{1}{2}$ mrad above or below the orbital plane, albeit at the expense of a drop in flux approaching one order of magnitude when compared with the orbital position.

Figure 1 shows a schematic diagram of the apparatus, which is elevated ~ 6 mm above the orbital plane. The beam is monochromated by a plane (220) germanium crystal and the 59-keV radiation is then scattered at 145° through an Fe (110) single crystal slice (stabilized with 3 wt% silicon) held between the poles of an electromagnet. The [111], [110] and [100] directions can be brought into parallelism with the magnetic field and the X-ray scattering vector simply by rotating the slice, and this novel transmission geometry is essential for such directional measurements. The Compton-scattered beam is detected by an intrinsic germanium detector and the data channelled into alternate memories according to the magnetic field direction.

In the absence of a magnetic field, the broadened spectrum of Compton-scattered radiation is related to the momentum distribution, $n(\mathbf{p})$, of all the target's electrons through the Compton profile integral $J(p_z) = \iint n(\mathbf{p}) dp_x dp_y$, z being

parallel to the scattering vector. The spin-dependent part, arising from the 3d and 4s electrons, can be isolated in an experiment with circularly polarized flux when the spins are flipped by the applied magnetic field. The difference spectrum can then be related to the magnetic Compton profile, $J_{\text{mag}}(p_z)$, defined as $J_{\text{mag}}(p_z) = \iint \sum_i [n_{i,\text{up}}(\mathbf{p}) - n_{i,\text{down}}(\mathbf{p})] dp_x dp_y$, where $n_i(\mathbf{p})$ is the momentum distribution in the i th 'up' or 'down' spin band. Each directional magnetic profile has an area equal to the number of electrons with unpaired spins, which is 2.2 per atom in iron; the experimental line shapes shown in Fig. 2 are therefore compared with theory after normalization to this value.

Although $J_{\text{mag}}(p_z)$ is still a projection of the spin density, it retains strong directional characteristics and provides a continuous function for comparison with theory, compared with the limited number of accessible data points from neutron diffraction studies. In the body-centred cubic crystal environment, neither the charge density nor the spin density is isotropic. Evidence for the former comes in diffraction studies from the inequality of 'paired' Bragg reflections occurring at the same momentum transfer¹⁸, and in Compton scattering experiments from the differences between directional Compton profiles¹⁹.

The spin density distribution, studied in earlier polarized neutron experiments^{20,21}, has been interpreted in terms of an excess of the majority spins along the [100] directions and a deficit (that is, a reversed net magnetization) in interatomic regions along [111] and [110] directions. This is qualitatively consistent with band calculations based on the linear combination of atomic orbitals (LCAO) (ref. 22) and the augmented plane wave (APW) (ref. 23) methods which link these features with the 3d and 4s-p bands respectively. The reversed magnetization density in regions far away from the atomic cores is associated with electrons having small momenta, and is therefore manifest as a decrease in the Compton profile around the origin, as shown in Fig. 2. This feature has been noted in previous studies of the polycrystalline profiles of iron^{12,13} and nickel¹⁴ and leads to a 'volcanic' line shape.

It is evident from the results shown in Fig. 2 that the 'crater' is greatest in the [111] profile and entirely absent from the [100] results. The APW calculation also shown in Fig. 2 is an extension of that cited earlier²³; it reproduces the observed behaviour for the [111] profile but overestimates the value at the origin for the [110] and [100] directions, the latter deviation being more marked. These differences are consistent with the earlier results on polycrystalline iron¹³⁻¹⁵, all of which implied a reversely magnetized component larger than predicted. These single-crystal results identify the crystal directions associated with the discrepancy.

It is not possible to transform the momentum-space results into the more familiar position-space densities, because charge densities in the two spaces are not directly related. Instead, it is necessary to consider the magnetic anisotropy in terms of the different momentum distributions assumed by electrons in the majority and minority spin Fermi distributions due to the spin-dependent potentials. The predicted anisotropic electron and hole Fermi surfaces indicate that, at low momentum, the minority spin contribution to the Compton profile which causes the dip is smallest when the projection direction, *z*, is aligned parallel to the [100] direction, and greatest for the [111] direction²⁴, in agreement with these results.

More elaborate investigations will be needed for a fuller understanding of the differences between experiment and theory reported here. For example, it may now be important to consider electron correlation effects and spin orbit interactions neglected in the present APW theory. With the improvements in statistical accuracy and resolution that should now evolve, magnetic Compton scattering experiments in anisotropic ferromagnetic metals will provide a critical test of band theories which may more than equal neutron scattering studies.

We thank the SERC for supporting this work and G. Clark of the Daresbury Laboratory and P. Collins of the Warwick University Physics Department for their assistance.

High-*T_c* superconducting phases in the series $\text{Bi}_{2.1}(\text{Ca}, \text{Sr})_{n+1}\text{Cu}_n\text{O}_{2n+4+\delta}$

J. L. Tallon, R. G. Buckley, P. W. Gilberd, M. R. Presland, I. W. M. Brown*, M. E. Bowden*, L. A. Christian & R. Goguel*

Physics and Engineering Laboratory and * Chemistry Division, Department of Scientific and Industrial Research, Private Bag, Lower Hutt, New Zealand

Following the report of Maeda *et al.*¹ on high-*T_c* superconductivity in samples of nominal composition $\text{BiSrCaCu}_2\text{O}_x$ we have identified three distinct phases in the homologous series $\text{Bi}_{2.1}(\text{Ca}, \text{Sr})_{n+1}\text{Cu}_n\text{O}_{2n+4+\delta}$ with *n* = 1, 2 and 3. These have zero resistance *T_s* of 80 K, 91 K and 105 K respectively. The structures in the homologous series appear to be based on alternating double bismuth Bi_2O_2 layers and perovskite $(\text{Sr}, \text{Ca})\text{O}-\text{CuO}_2$ layers, with higher members obtained by intercalating additional 3.1 Å $\text{Ca}-\text{CuO}_2$ bilayers. The structures can be indexed on a pseudotetragonal subcell with *a* = *b* = 5.4 Å and *c* = 24.4 Å (*n* = 1), 30.76 Å (*n* = 2) or 36 Å (*n* = 3). We were unable to prepare the *n* = 4 member. In each structure a 2× superlattice structure in the *c*-direction arises as a natural consequence of a 19/4 incommensurate structure in the *b* direction, which accounts for 2.1 Bi atoms in the unit formula. The critical temperature, *T_c*, for these phases is sharply dependent on the Sr/Ca ratio and oxygen stoichiometry, as determined by heat treatment.

The $\text{Bi}_2\text{O}_3-\text{CaO}-\text{SrO}-\text{CuO}$ phase diagram was explored by preparing numerous compositions and reacting and sintering at temperatures ranging from 830 °C to 870 °C in a variety of atmospheres. From these studies, together with electron microscopy and X-ray powder diffraction, we have been able to identify the conditions under which the three superconducting phases can be synthesized. We have subsequently learned of the identification of similar phases in the $\text{Tl}-\text{Ba}-\text{Ca}-\text{Cu}-\text{O}$ system²⁻⁴, and of the observation of intergrowths of these phases in $\text{BiCaSrCu}_2\text{O}_x$ ^{5,6}. Following rapidly established convention, we refer to the three phases discussed here ($\text{Bi}_{2.1}\text{CaSrCuO}_{4+\delta}$, $\text{Bi}_{2.1}\text{CaSr}_2\text{Cu}_2\text{O}_{8+\delta}$ and $\text{Bi}_{2.1}\text{Ca}_2\text{Sr}_2\text{Cu}_3\text{O}_{10+\delta}$) as '2111', '2122' and '2223' respectively.

Samples with starting compositions near 2122 and with a range of strontium-to-calcium ratios were prepared under a variety of reaction, sintering, annealing and quenching regimes. The following conclusions may be drawn: (1) *T_c* is maximized

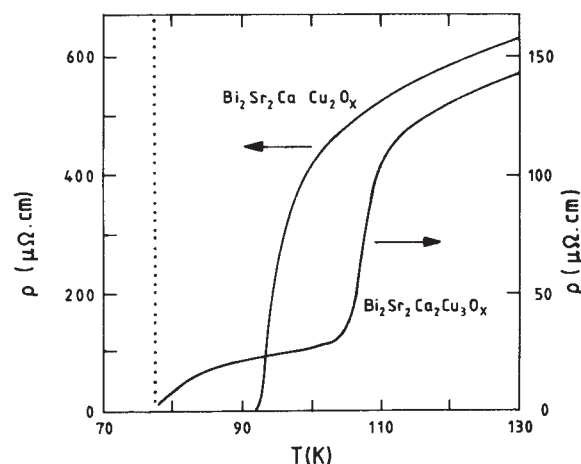


Fig. 1 Resistivity as a function of temperature for the 91 K superconductor $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_x$ and a mixed-phase ceramic containing ~15% of the 105 K superconductor $\text{Bi}_2\text{Ca}_2\text{Sr}_2\text{Cu}_3\text{O}_x$.

Received 28 March; accepted 20 April 1988.

1. Lipps, F. W. & Tolhoek, H. A. *Physica* **20**, 395-405 (1954).
2. Blume, M. *J. appl. Phys.* **57**, 3615-3618 (1985).
3. Lovesey, S. W. *J. Phys. C* **20**, 5625-5639 (1987).
4. Platzman, P. & Tzozar, N. *Phys. Rev.* **B2**, 3556-3559 (1970).
5. deBergerin, F. & Brunel, M. *Phys. Lett.* **A39**, 141-142 (1972).
6. deBergerin, F. & Brunel, M. *Acta crystallogr.* **A37**, 314-324 (1981).
7. Brunel, M. & deBergerin, F. *Acta crystallogr.* **A37**, 324-331 (1981).
8. Blume, M. & Gibbs, D. *Phys. Rev.* **B37**, 1779-1789 (1988).
9. Gibbs, D., Moncton, D. E. & D'Amico, K. L. *J. appl. Phys.* **57**, 3619-3622 (1985).
10. Vettier, C. *et al. Phys. Rev. Lett.* **56**, 757-760 (1986).
11. Cooper, M. J. *Rep. Prog. Phys.* **48**, 415-418 (1985).
12. Holt, R. S. & Cooper, M. J. *Nucl. Instrum. Meth.* **213**, 447-452 (1983).
13. Sakai, N. & Sekizawa, H. *Phys. Rev.* **B36**, 2164-2169 (1987).
14. Mills, D. M. *Phys. Rev.* **B36**, 6178-6181 (1987).
15. Cooper, M. J. *et al. Phys. Rev.* **B34**, 5984-5987 (1986).
16. Brunel, M., Patrat, G., deBergerin, F., Rousseaux, F. & Lemmonier, M. *Acta Crystallogr.* **A39**, 84-88 (1983).
17. Schütz, G. *et al. Phys. Rev. Lett.* **58**, 737-740 (1980).
18. Diana, M. & Mazzone, G. *Phys. Rev.* **B9**, 3898-3904 (1974).
19. Rollason, A. J., Holt, R. S. & Cooper, M. J. *J. Phys. F* **13**, 1807-1819 (1983).
20. Schull, C. G. & Yamada, Y. *Phys. Soc. Jap. Suppl.* **B17**, 1 (1962).
21. Schull, C. G. & Mook, H. A. *Phys. Rev. Lett.* **16**, 184-186 (1966).
22. Tawil, R. A. & Callaway, J. *Phys. Rev.* **B7**, 4242-4252 (1973).
23. Wakoh, S. & Kubo, Y. *J. magn. magn. Mater.* **5**, 202-211 (1977).
24. Rollason, A. J. thesis, Univ. Warwick (1984).

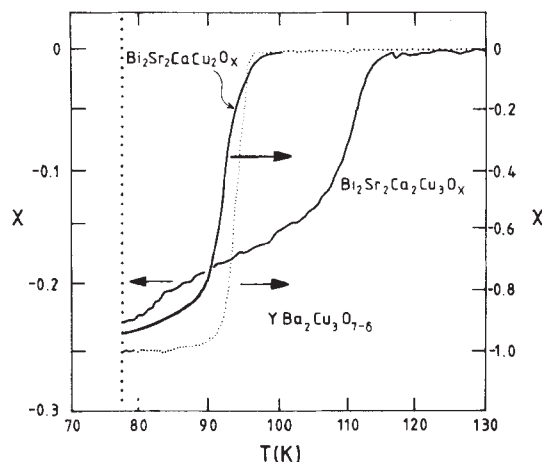


Fig. 2 Temperature-dependence of the a.c. susceptibility of a 2122 and a mixed-phase sample containing 2223 normalized to the saturated susceptibility of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ (also shown).

by reacting and sintering in the partial melting region at 860–870 °C, followed by a rapid quench. The resistivity plot shown in Fig. 1 for a single-phase 2122 sample so prepared exhibits zero resistance at 91 K. Without the rapid quench, T_c falls to ~80 K or lower, and this is the state reported by others^{7–9}.

Thermal gravimetry and high-temperature resistivity studies (R.G.B., J.L.T., N. Milestone and A. Bittar, unpublished) show that the elevation of T_c by quenching is due to oxygen deficiency. Annealing at lower temperature increases oxygen content, depresses T_c and, surprisingly, reduces the bulk resistivity. (2) Bulk superconductivity is confirmed by the a.c. susceptibility measurements shown in Fig. 2. The diamagnetic signal is normalized to that for a standard $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ sample and it reaches 90% of the fully saturated $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ signal at 77 K. (3) Platelets of dimension 1 mm × 1 mm × 0.1 mm removed from samples with $T_c \sim 91$ K were magnetically levitated in liquid oxygen (boiling point 90 K) and these showed compositions very close to $\text{Bi}_{2.1}\text{CaSr}_2\text{Cu}_2\text{O}_x$. Within multi-phase samples, a spread in the Sr-to-Ca ratio ($\pm 15\%$) and the Cu-to-Bi ratio ($\pm 5\%$) was measured, of which either could account for an observed broadening of the X-ray diffraction peaks and the regular occurrence of a 'toe' in the resistivity drop at T_c . These variations probably arise from the propensity of this material for intergrowth of the $n = 1, 2$ and 3 phases^{5,6}.

Others^{7–10} have investigated the structure of $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_x$. From X-ray powder diffraction we also observe a structure in which most reflections can be indexed with a pseudo-tetragonal sub-cell with $a \approx b = 5.4$ Å and $c = 30.76$ Å. The particular values depend on the precise Sr-to-Ca ratio, with variations of 0.5% in the c -direction and 0.2% in the a - and b -directions being observed. The structure is a -centred, comprising two

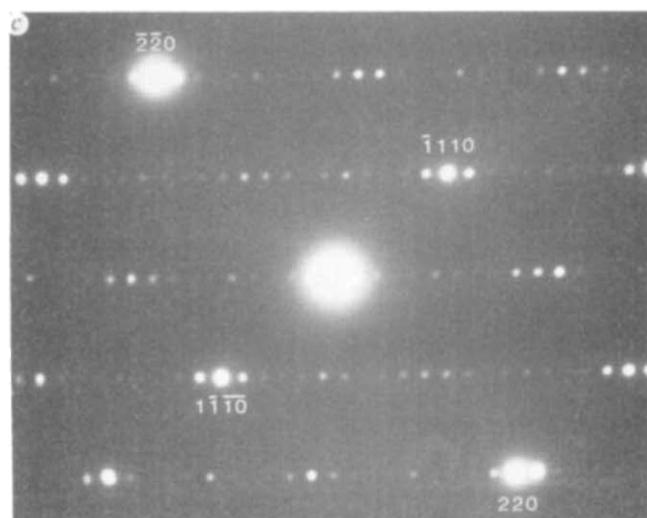
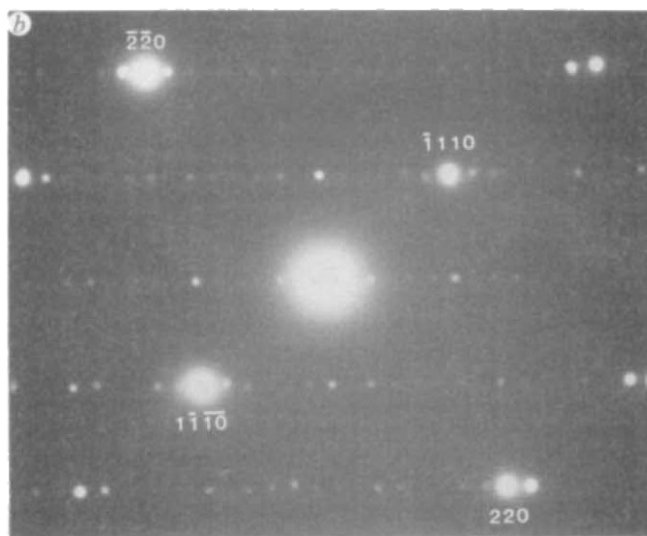
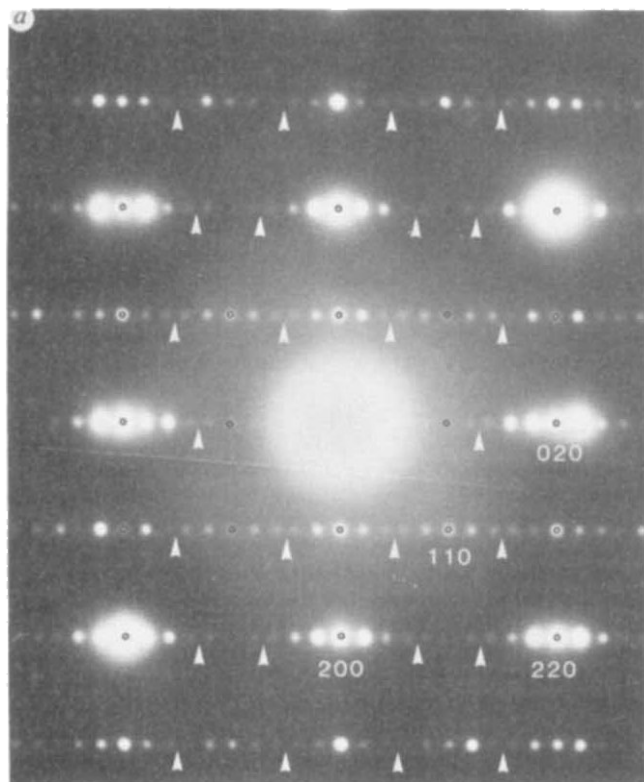


Fig. 3 *a*, The [001] zone axis electron diffraction pattern for 2122 showing the 25.7 Å incommensurate structure in the b -direction. The spots representing the 5.4 Å × 5.4 Å unit cell are marked. Main reflections (020, 220 and so forth) are flanked by two or three satellites on each side, whereas weak reflections have only one or two. The effect is to create occasional closely spaced spots (arrowed), proving that their origin is an incommensurate structure rather than a superlattice of the basic 5.4 Å cell. *b*, The [551] zone axis pattern for 2122, indexed on a $5.4 \times 5.4 \times 61.4$ Å³ unit cell. Taking the periodicity of the incommensurate structure in the b -direction to the 25.7 Å (from Fig. 3*a*), the spacing of the satellites along the [015] reciprocal lattice vector can only be fitted with a unit cell for which $c = 61.4$ Å. *c*, The [551] zone axis pattern for 2223, indexed on a $5.4 \times 5.4 \times 72$ Å³ unit cell.

15.4 Å CuO₂-SrO-BiO-BiO-SrO-CuO₂-Ca-sequences^{7,9} mirrored about the Ca. However, selected-area electron diffraction is more sensitive to superstructure and clearly establishes at 77 K the following additional features, as illustrated in Fig. 3: (1) in the *a*-direction, a cell dimension of 5.4 Å; (2) in the *b*-direction, a cell dimension of 5.4 Å, and a 19/4× incommensurate structure of 25.7 Å; (3) in the [551] zone axis diffraction pattern, spots are observed that are only consistent with a *c*-axis dimension of 61.4 Å, that is, there is a 2× *c*-axis superstructure.

As shown in Fig. 4, this *c*-axis superstructure has a simple explanation in the nature of the 19/4 incommensurate structure in the *b*-direction. Attributing this 25.7 Å structure to a grouping of five Bi₂O₂ subunits within a Bi layer, the results can be interpreted in terms of 19 Cu pairs being commensurate with 20 Bi pairs, giving a true *b*-axis superlattice of $4 \times 25.7 = 102.8$ Å. For every five Bi pairs, the mismatch advances by 1/4. If the stacking of the Bi bilayers in the *c*-direction were such as to stagger the incommensurate structure in the *b*-*c* plane by a phase lag of 1/4 for each unit, then the structure would repeat every four bilayers, the *c*-axis being doubled to 61.4 Å. This means that the Bi₂O₂ sublattice is monoclinic. Further evidence for this interpretation lies in the observation from electron beam X-ray analyses of Bi:Cu ratios in single crystals of 2.1:2.0, rather than 2:2.

No evidence is seen for the 5× superstructure in the *b*-direction reported by others^{7,8}. From high-resolution images, Hewat *et al.*¹⁰ report a variable blocking of predominantly 5×, but sometimes 4× and 6×, rather than the uniform 19/4× structure that we propose. We can only note that their samples were prepared from non-stoichiometric compositions at lower temperatures, and therefore may have different cation and anion stoichiometries, as illustrated by the much lower zero resistance *T_c* of 60 K (ref. 9).

Attempts to prepare the third member of the series 2223 from stoichiometric starting materials have not been wholly successful, due to phase separation to 2122, (Sr, Ca)CuO₂ and only small amounts of 2223. The phase can be prepared at a 10 to 20% level from a 1112 or a 1113 mixture by reacting and sintering at 860 °C, followed by a rapid quench. Unlike 2122, this phase requires oxygen loading to maximize *T_c* and also is sensitive to ambient moisture. The effect of hour-long anneals in air, followed by quenching into liquid nitrogen from progressively lower temperatures, was to successively lower the normal state resistivity and displace *T_c* to higher temperatures. This process is optimized at ~500 °C, producing materials with *T_c* ~ 105 K. Zero resistance can be achieved at 85 K in this system, but this has to be compromised to promote the 105 K transition. Both *T_c* and the normal-state conductivity are sharply maximized at a strontium-to-calcium ratio of 1:1. The resistivity curve for a sample annealed at 400 °C in air is shown in Fig. 1, whereas Fig. 2 includes the a.c. susceptibility of this sample. Diamagnetism sets in at ~115 K, and by 77 K the sample containing ~15% 2223 shows a diamagnetic signal that is 23% of that of the fully saturated YBa₂Cu₃O_{7-δ}. Like the 2122 material, crystals of 2223 are platy and exhibit a 5.4 Å × 5.4 Å cell in the basal plane, with the same 19/4 incommensurate superlattice in the *b*-direction. Diffraction patterns for the [551] zone axis shown in Fig. 3c can be indexed on a 5.4 × 5.4 × 72 Å cell. X-ray powder diffraction showed a predominance of 2122, with an additional broad basal reflection corresponding to a *c*-repeat of ~18 Å. This leads to the natural conclusion that Bi₂Ca₂Sr₂Cu₃O_x is structurally related to Bi₂CaSr₂Cu₂O_x by the insertion of an additional pair of Ca-CuO₂ sheets. The thickness of this pair of sheets in 2122 is ~3.1 Å (ref. 9), so the increase in stacking periodicity from 15.4 Å in 2122 to 18 Å in 2223 is reasonable.

The structurally pure first member of the homologous series discussed here is 2021, and this was evidently the active phase in the previously reported material^{6,11} Bi₂Sr₂Cu₂O_{7+δ} with *T_c* ≤ 22 K and *a* = 5.22 Å, *b* = 26.6 Å and *c* = 48.8 Å. Incidentally,

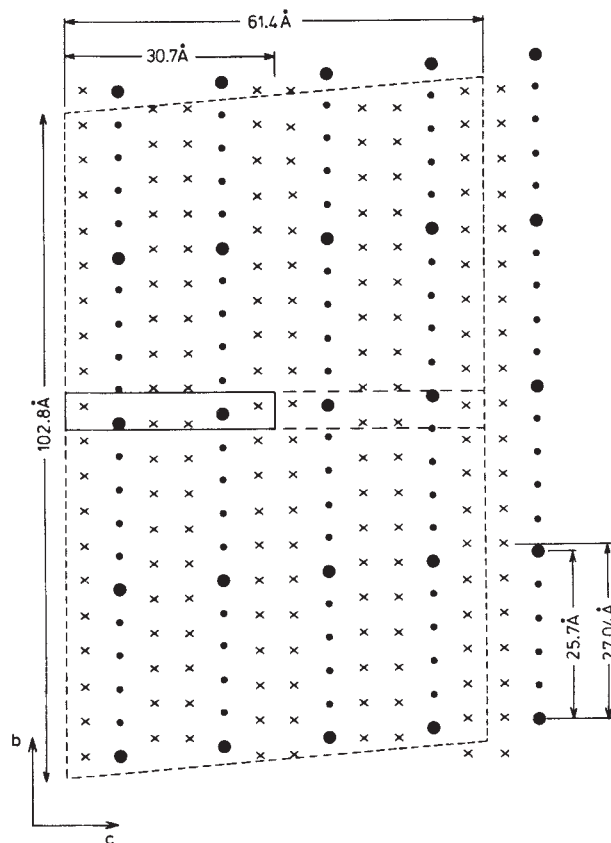


Fig. 4 Schematic diagram of the proposed superstructure in the *b*-*c* plane of Bi_{40/19}CaSr₂Cu₂O_{8+δ}, where the dots are representative of two Bi pairs, one from each layer of the Bi bilayer, and the crosses denote single Cu pairs. Strontium, calcium and oxygen atoms are omitted for clarity. Every fifth dot in the *b*-direction is enlarged to assist viewing of the structure. The phase repeats in the *b*-direction every 19 Cu pairs (or every 20 Bi pairs within a single Bi layer) giving a *b*-supercell length of 102.8 Å. The 5.4 Å × 30.7 Å subcell is indicated by the continuous line box, the 5.4 Å × 61.4 Å cell by the continuous plus dashed line boxes and the true 102.8 Å × 61.4 Å supercell by the short-dashed box. Every consecutive Bi bilayer is staggered in the *b*-direction by one quarter of the Cu pair-Cu pair distance, and thus the phase repeats in the *c*-direction every 4 bilayers to give *c* = 61.4 Å.

this *c*-axis dimension supports the doubling of the *c*-axis in this series proposed above. We find an additional related member 2111 with *T_c* ~ 80 K, where the strontiums and calciums are presumably intermixed. Reacting from stoichiometric proportions at 815 °C resulted in multi-phase samples with nearly equal quantities of 2111 and 2122, and also some 2223. Two resistive and magnetic transitions identifiable with the main superconducting components are observed giving a zero-resistance *T_c* ~ 80 K for 2111. This phase can be indexed with the same structure in the *a*-*b* plane as 2122 and a *c*-axis of 48.8 Å, suggesting a structure comprising BiO-BiO-(Sr, Ca)O-CuO₂-(Sr, Ca)O-repeats.

These bismuth-based superconductors have rather lower resistivities than YBa₂Cu₃O_{7-δ}. Resistivities of 50 μΩ cm just before superconducting onset are frequent, but one sample prepared from a 2223 composition, although comprising predominantly 2122, exhibited resistivities at room temperature of 30 μΩ cm and, at onset, of 7 μΩ cm. This is the lowest normal-state resistivity we have observed in the wider class of oxide superconductors.

In summary, we have identified Bi_{2.1}CaSrCuO_x (2111), Bi_{2.1}CaSr₂CaCu₂O_x (2122) and Bi_{2.1}Ca₂Sr₂Cu₃O_x (2223) as superconducting phases in a homologous series with optimized zero resistance at 80 K, 91 K and 105 K (extrapolated) with

onsets at 115 K. Normal-state and superconducting properties of these phases are dependent on oxygen stoichiometry. Rapid quenching of 2122 to an oxygen-deficient state advances T_c from ~ 80 K up to 91 K. In contrast, for 2223, oxygen loading by annealing at 500 °C is required to maximize T_c . Contrary to previous structural studies, we believe the 2122 to be pseudotetragonal with a $5.4 \times 5.4 \times 30.76$ Å³ subcell and a $2 \times c$ -axis superlattice which arises from the staggering of a $19/4 \times$ incommensurate superlattice in the b -direction. This yields lattice constants of $a = 5.4$ Å, $b = 102.8$ Å and $c = 61.5$ Å. The structures of the 2111 and 2223 phases are similar with additional ~ 3.1 Å Ca-CuO₂ sheets respectively removed or added. This yields a stacking periodicity in the c -direction of 12.2 Å and 18 Å respectively, and c -axes of 48.8 Å and 72 Å.

We are grateful to A. H. Schuitema for technical assistance, K. Card and K. Ryan for assistance with electron microscopy and D. Rogers for differential thermal analysis measurements.

Received 7 April; accepted 18 April 1988.

1. Maeda, H., Tanaka, Y., Fukutomi, M. & Asano, T. *Jap. J. appl. Phys. Lett.* (in the press).
2. Sheng, Z. Z. & Hermann, A. M. *Nature* **332**, 138–139 (1988).
3. Ginley, D. S. *et al. Physica C* (in the press).
4. Parkin, S. S. P. *et al. Phys. Rev. Lett.* (in the press).
5. Ikeda, S. *et al. Jpn. J. appl. Phys. Lett.* (in the press).
6. Morgan, P. E. D., Ratto, J. J., Housley, R. M. & Porter, J. R. *MRS Spring Symp., Reno, NV* (Materials Research Society, Pittsburgh, 1988).
7. Subramanian, M. A. *et al. Science* **239**, 1015–1017 (1988).
8. Hazen, R. M. *et al. Phys. Rev. Lett.* **60**, 1174–1176 (1988).
9. Bordet, P. *et al. Proc. Interlaken Conf. on High T_c Superconductors Physica C* (in the press).
10. Hewat, E. A. *et al. Nature* (in the press).
11. Michel, C. *et al. Z. Phys.* **B68**, 421–423 (1987).

New evidence from the South China Sea for an abrupt termination of the last glacial period

W. S. Broecker*, M. Andree†, M. Klas*, G. Bonani‡, W. Wolfli‡ & H. Oeschger†

* Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA

† Physics Institute, Universität Bern, Sidlerstrasse 5, Bern, Switzerland

‡ Laboratorium für Kernphysik, ETH Honggerberg, 8093 Zurich, Switzerland

Abrupt changes in climatic conditions have been seen at high latitudes in the North Atlantic¹ and the Antarctic^{2,3} at 13 kyr BP. It is important to determine whether this abrupt change was confined to high-latitude regions or whether it was global. Here we present results demonstrating an abrupt change in the rate and character of sedimentation in the South China Sea at the close of the last glacial period. Radiocarbon dating and its position in the oxygen isotope shift suggest that this change may be coincident with the changes found at high latitudes.

In their classic study of sediments from the northern Atlantic, Ruddiman *et al.*¹ demonstrated that the warming of surface waters associated with the end of the last glacial (termination I) and with penultimate glaciation (termination II) occurred very rapidly. A radiocarbon chronology for one of these records (core V23-81) suggests that the pronounced warming of surface waters in this area began about ≈ 15 kyr BP and was completed by ≈ 13 kyr BP². Large-scale melting of the Laurentian ice sheet began ≈ 13.5 kyr BP years ago, as shown by radiocarbon dating of the onset of the melt-water event in the Gulf of Mexico⁴. Although not yet dated directly by radiocarbon, a sharp change in surface water conditions is also found in ice cores from the Antarctic^{2,3}. For both terminations I and II the sudden demise of *Cycladophora davisiana*, which heralds a change in surface-

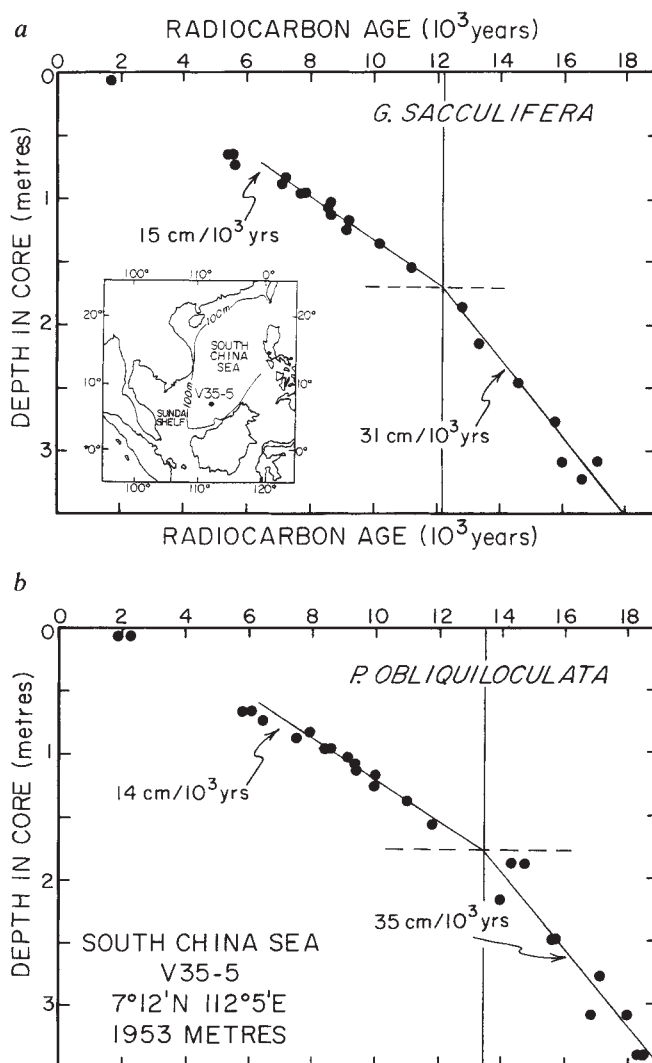


Fig. 1 Plot of radiocarbon age against depth in South China Sea core V35-5: *a*, *G. sacculifera*; *b*, *P. obliquiloculata*. The dashed lines designate the depth of the midpoint of the abundance change of planktonic foraminifera (see Fig. 2). As surface ocean water ΣCO_2 has a lower $^{14}\text{C}/^{12}\text{C}$ ratio than atmospheric CO_2 , 400 yr has been subtracted from these ages so they can be compared with those organic matter formed from atmospheric CO_2 .

water conditions, occurs near the beginning of the oxygen isotope shift. Correlation with radiocarbon-dated oxygen isotope records⁷ suggests an age of between 13 and 14 kyr BP for this event.

It should be pointed out that a second core from the northern Atlantic has been radiocarbon dated by Bard *et al.*⁵ who obtained an age of 11.5 kyr BP for the onset of warm conditions off the British Isles. While we envision a simultaneous change for the entire northern Atlantic region, Bard *et al.* interpret their results to indicate a gradual migration of the front separating subpolar and polar waters. They contend that this migration began ≈ 13 kyr ago off southern Portugal and did not reach Ireland until 11.5 kyr BP⁶. The only explanation we can see for the 1,500-yr difference in the chronologies for two nearby cores is that a hiatus exists in the core studied by the French. For this reason we stress the results from core V23-81.

We have made a detailed study of a core from the South China Sea (V35-5). It comes from a knoll (depth 1,953 m) in the southern part of the basin (7°11' N, 112°5' E). The chronology for this core is documented by radiocarbon measurements made

by accelerator mass spectrometry on hand-picked planktonic shells⁸. Details of the measurement procedures are published elsewhere^{9,10}. A report containing a listing of all our accelerator radiocarbon measurements is available on request¹¹. Separate measurements were made on shells of the species *Globigerinoides sacculifera* and *Pulleniatina obliquiloculata*. Faunal abundances (mg shell per gram of sediment) were determined as part of the radiocarbon measurement program. An oxygen isotope record for benthic foraminiferous from this core has been published¹².

Table 1 lists the results; Fig. 1 shows plots of radiocarbon age against depth in the core for the results on *G. sacculifera* and *P. obliquiloculata*. The sedimentation rate for Holocene time is $\approx 15 \text{ cm kyr}^{-1}$ while that for glacial time is $\approx 33 \text{ cm kyr}^{-1}$. If best-fit line segments are drawn through the Holocene points and through the glacial points, they intersect at a depth of $\approx 175 \text{ cm}$.

Analyses of samples from the 70–75 cm (Holocene) and 230–234 cm (peak glacial) levels in this core have been carried out to determine whether a different source existed for the excess glacial material (F. Grousset, personal communication). The silicate fractions from these two horizons were found to be remarkably similar in mineralogy and major element chemistry. They were also found to be identical in neodymium isotope composition. Uranium and thorium isotope measurements on these two samples have been carried out and no significant difference has been found (Y. Lao, personal communication). Both had, as expected, only a very small excess of ^{230}Th (over ^{234}U) and ratios of ^{234}U to ^{238}U of ~ 1 . Hence a distant source (for example, loess from China) need not be called upon to explain the high glacial accumulation rate.

The abundances of both *G. sacculifera* and *P. obliquiloculata* undergo abrupt increases from glacial and Holocene time (see Fig. 2). A similar change is seen for the total abundance of planktonic foraminifera shells. These changes are centred at a depth of 175 cm and occur over a depth interval $\leq 25 \text{ cm}$. The tenfold increase in the abundance of *P. obliquiloculata* between glacial and Holocene time cannot be explained by the doubling of sedimentation rate alone; in addition a fivefold increase in the accumulation rate of *P. obliquiloculata* must have occurred. A similar increase occurs in the total foraminifera shell abundance. Thus the South China Sea region experienced an abrupt climate change, which resulted in both a change in the rate of sediment accumulation and a change in the productivity of planktonic foraminifera.

Radiocarbon dating of the transition has been done by analys-

Table 1 Abundance and radiocarbon ages for *G. sacculifera* and *P. obliquiloculata* from South China Sea core V35-5.

Depth (cm)	Abund. <i>G. sac.</i> (mg g ⁻¹)	¹⁴ C age <i>G. sac.</i> (yr)	Abund. <i>P. obliq.</i> (mg g ⁻¹)	¹⁴ C age <i>P. obliq.</i> (yr)
135–140	0.84	10,490 $\pm$ 130	—	—
140–145	1.54	10,800 $\pm$ 120	—	—
130–145	—	—	1.24	11,010 $\pm$ 190
150–160	1.19	11,180 $\pm$ 200	1.43	11,810 $\pm$ 190
160–165	0.72	—	1.35	12,520 $\pm$ 210
160–165	0.61	—	1.20	—
165–170	0.54	—	0.84	—
170–175	0.45	11,460 $\pm$ 190	0.77	12,770 $\pm$ 210
170–175	0.48	—	0.86	—
175–180	0.22	—	0.43	—
180–195	0.32	12,840 $\pm$ 190	0.17	14,760 $\pm$ 220
180–195	0.27	12,820 $\pm$ 190	0.16	14,380 $\pm$ 210
205–220	0.15	—	0.07	—
210–220	0.23	13,340 $\pm$ 190	0.05	13,940 $\pm$ 200

The radiocarbon ages have been corrected for the atmosphere–surface-ocean radiocarbon difference (that is, 400 years have been subtracted from the conventional radiocarbon ages). The dashed line denotes the mid-depth of the abundance change.

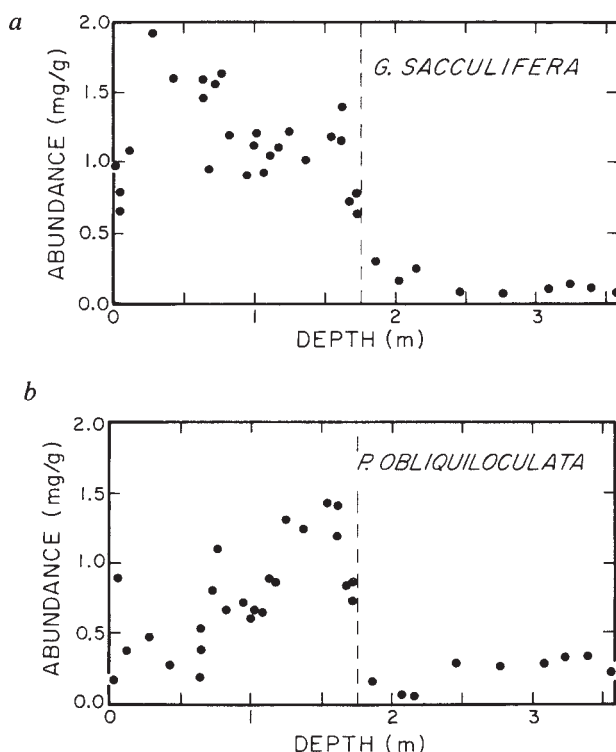


Fig. 2 Abundance of the shells of two planktonic species of foraminifera (a, *G. Sacculifera*; b, *P. obliquiloculata*) as a function of depth in South China Sea core V35-5. The shells were hand-picked from the $>150 \mu\text{m}$ size fraction.

ing separately hand-picked samples of *G. sacculifera* and *P. obliquiloculata*. Unfortunately, the ages for these two species differ significantly (see Table 1). As Fig. 2 shows, the *G. sacculifera* results yield an age of 12.2 kyr for the 175-cm level whereas the *P. obliquiloculata* results yield an age of about ≈ 13.5 kyr. In a separate paper¹³ we explore possible reasons for the anomalies in these ages but do not come up with any satisfactory answer. Thus at this point we cannot say which of the ages is to be preferred.

The oxygen-isotope record for benthic foraminiferous¹² shows a gradual decrease over the depth interval 180–70 cm (see Fig. 3). This record is consistent with the generally held interpretation that the ice caps of the last glaciation began melting ≈ 15 kyr ago and disappeared ≈ 6 kyr ago¹⁴. The position in the oxygen-isotope record of the abrupt change in the abundance of planktonic foraminifera shells is nearly the same as the position of the *C. davisiana* change in the oxygen isotope record for Antarctic core RC11–120 (ref. 3).

We may speculate about the cause of the change in sedimentation rate in the South China Sea. The most likely source of the detritus, which makes up the sediment, is rivers flowing from Southeast Asia, Malaysia, Sumatra and Borneo onto the vast Sunda Shelf. During glacial time this shelf was laid bare by the lowered sea level, and so the rivers must have crossed the shelf and entered the South China much closer to the present site of core V35-5. Thus one possible explanation for the sedimentation rate change is that as sea level rose at the end of glacial time, it flooded the Sunda Shelf causing the sediment delivered from the adjacent land masses to accumulate on the shelf instead of in the South China Sea. The timing of the change, however, casts doubt on this explanation. It occurred close to the beginning of the benthic oxygen isotope decrease (and hence also of the sea level rise). Because the water depth on the Sunda Shelf is everywhere $<100 \text{ m}$ in depth, its inundation by the rising sea probably did not even begin until after 13 kyr BP.

Another possibility is that the climate change that terminated glacial conditions in this area brought about a major change in the vegetative cover. Most of the area is now tropical rainforest. If, as has been claimed¹⁵, drier conditions prevailed in the Southeast Asia region during glacial time, then the reduced vegetative cover (savannah rather than rainforest) would have less effectively held back soil erosion. Under such a circumstance, despite the lower rainfall, glacial erosion rates would have been higher. This is our preferred explanation. Unfortunately, the pollen records for this region are sparse, and according to Flenley¹⁶ those that do exist contain no conclusive evidence for drier conditions.

The increase in the accumulation rates of *G. sacculifera* and *P. obliquiloculata* shells could reflect either a change in the productivity of these organisms in the overlying surface waters or a decrease in the extent of post-depositional dissolution and breakup. As the state of preservation of both shell types in the glacial section of the core is quite good we reject the second possibility, and interpret the abundance changes as indication of a major modification in the ecological situation in South China Sea surface waters.

Our approach of using the absolute abundance of the shells of single species of planktonic Foraminifera (that is, weight of shells per weight of bulk sediment) as an indicator of changing surface ocean conditions constitutes a departure from the norm in palaeoceanography. Nearly all such studies involve the use of relative abundances (that is, the ratio of the number of shells of a given species to the total number of planktonic shells). The objective of these more traditional studies is to obtain temperature histories of surface waters. For example, a study of cores from the Sulu Sea (adjacent to the South China Sea) reveals that the relative abundances of the planktonic species *Neogloboquadrina pachyderma* and *P. obliquiloculata* underwent changes at the time of the onset of the oxygen isotope change¹⁷.

The abruptness of the close of glacial time has an important implication for theories regarding the cause of cyclic glaciation. A strong case has been made that the changes in seasonality induced by cyclic changes in the Earth's orbital parameters drive glacial cycles^{18,19}. The most commonly invoked linkage between seasonality and climate has been the nonlinearity of the response of ice cover²⁰⁻²², in that during times of enhanced seasonality, excess summer melting more than compensated for the impacts of the more intense winters. Were this the linkage, then climate should change in accord with the extent of ice cover. Faunal data from ocean cores suggests that this is not so: the cold period came to an abrupt close; the ice then melted in response to this change.

Hence if the Milankovitch hypothesis of orbital forcing is correct, some other linkage must be postulated. Broecker, *et al.*²³ have suggested that the linkage involves reorganizations of the manner in which the ocean-atmosphere operates, and point to the transport of water vapour through the atmosphere from the Atlantic to the Pacific as an important constraint on the pattern of large-scale ocean circulation. The reason is that this transport leads to a buildup of the salt content of waters in the Atlantic. The ocean compensates by establishing a current system that trades the more salty waters of one basin for the less salty waters of the other basin.

Broecker *et al.*²³ postulate further that the low seasonal contrast in glacial time caused the ocean to operate in a different mode from today's. In this alternative mode, the Atlantic conveyor belt was turned off (or at least down) cutting off the supply of ocean heat to the atmosphere over the northern Atlantic. Then, as the contrast in Northern Hemisphere seasonality increased toward its maximum, the system flipped back into its current mode of operation. Except for a brief interval during the Younger Dryas event, the system has remained locked in its current conveyor mode through the past 13 kyr.

No evidence for the Younger Dryas event (≈ 10.8 – 10.0 kyr BP), which appears so prominently in cores from the

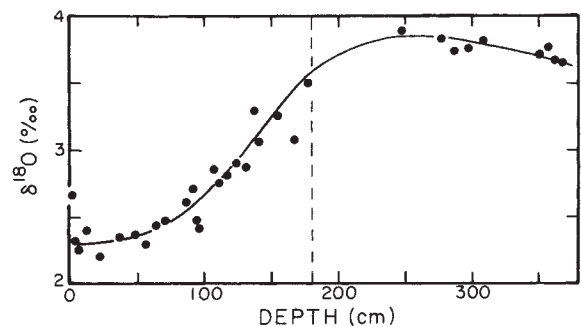


Fig. 3 Oxygen isotope record for benthic foraminifera from core V35-5 (12). The dashed line is the depth in the core at which the abrupt change in foraminiferal abundance is centred.

northern Atlantic^{3,24}, can be seen in the record from the record for V35-5. Were the brief cold episode a global return to the climate of glacial time, it should be seen in the abundance records of planktonic foraminifera. Figure 3 shows that no anomaly appears in either the *G. sacculifera* or the *P. obliquiloculata* abundance record in the depth interval 125–140 cm, which corresponds to the Younger Dryas time interval. This is consistent with the hypothesis that the Younger Dryas was caused by a brief shut down of the northern Atlantic conveyor resulting from a sudden diversion of the discharge of meltwater from the Mississippi to St Lawrence drainage¹³. While causing a drastic cooling in the northern Atlantic region²³ this shut down did not persist long enough to permit the global mode of ocean-atmosphere operation to be re-established.

Combined with evidence from the northern Atlantic region and from the Antarctic, we suggest that a major global change in climate occurred between 14 and 13 kyr BP. The abruptness of this change suggests that if orbital cycles are the pacemaker of the ice ages, then the response to this forcing must be of the monsoon variety. Rather than changing gradually with seasonality, the Earth's climate system responds through reorganizations of the ocean's circulation pattern. We postulate that the link between seasonality and ocean circulation involves the coupling between the transport of water vapour through the atmosphere and salt through the sea.

We thank Dorothy Peteet for helpful discussions about the pollen records from Southeast Asia. Financial support for this work was provided by NSF Climate Dynamics Program, by a NASA grant, and by the Swiss National Science Foundation.

Received 1 September 1987; accepted 18 February 1988.

- Ruddiman, W. F., Sancetta, C. D. & McIntyre, A. *Phil. Trans. R. Soc. B*, **280**, 1191–142 (1977).
- Hays, J. D., Lozano, J. A., Shackleton, N. & Irving, G. *Mem. geol. Soc. Am.* **145**, 337–372 (1976).
- Cooke, D. W. & Hays, J. D. *Proc. 3rd Symp. Antarct. Geol. Geophys.* (ed. Craddock, C.) 1017–1025 (University of Wisconsin Press, Madison, 1982).
- Broecker, W. S. *et al. Paleoceanography* **3**, 1–19 (1987).
- Bard, E. *et al. Climate Dynam.* **1**, 101–112 (1987).
- Bard, E. *et al. Nature* **328**, 791–794 (1987).
- Broecker, W. S. *et al. Quat. Res.* (in press).
- Andree, M. *et al. Climate Dynam.* **1**, 53–62 (1986).
- Andree, M. *et al. Nucl. Instr. Meth. B5*, 274–279 (1984).
- Suter, M. *et al. Nucl. Instr. Meth. B5*, 117–122 (1984).
- Broecker, W. S. *et al. AMS ¹⁴C Data List Rep. no. 1* (available from M. Klas, LDGO, Palisades, New York) (1987).
- Oppo, D. W. & Fairbanks, R. G. *Earth-planet. Sci. Lett.* **86**, 1–15 (1987).
- Broecker, W. S. *et al. Paleoceanography* (submitted).
- Mix, A. C. & Ruddiman, W. F. *Quat. Sci. Rev.* **4**, 59–108 (1985).
- Verstappen, H. T. *Geol. J.* **4**, 45–54 (1980).
- Flenley, J. *Equatorial Rain Forest: Geological History* 77–100 (Butterworths, London, 1978).
- Linsley, B. K. *et al. Mar. Micropaleontol.* **9**, 395–418 (1985).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
- Imbrie, J. *et al. in Milankovitch and Climate Part I* (eds Berger, A. L. *et al.*) 269–305 (Reidel, Dordrecht, 1984).
- Milankovitch, M. in *Handbuch der Klimatologie I(A)* (eds Koppen, W. & R. Geiger) 1–176 (Borntraeger, Berlin, 1930).
- Weertman, J. *Nature* **261**, 17–20 (1976).
- Imbrie, J. & Imbrie, J. Z. *Science* **207**, 943–953 (1980).
- Broecker, W. S., Peteet, D. & Rind, D. *Nature* **315**, 21–25 (1985).
- Duplessy, J. C. *et al. Nature* **320**, 350–352 (1986).

Vehicles and aircraft on floating ice

V. A. Squire*, W. H. Robinson†, P. J. Langhorne‡
& T. G. Haskell§

Department of *Mathematics and Statistics, and ‡Department of Physics, University of Otago, PO Box 56, Dunedin, New Zealand
† Physics and Engineering Laboratory, and §Division of Information Technology, Department of Scientific and Industrial Research, Private Bag, Lower Hutt, New Zealand

Ice roads and ice runways are a common feature of Arctic and Antarctic transportation: cars are driven across floating ice, whether for pleasure or for economic reasons such as avoiding a circuitous route around a bay or lake; logging trucks take short cuts from forest to mill; and aircraft routinely land and take off. Occasionally, catastrophic breakthrough and consequent injury or loss of life occurs. Although theoretical work to calculate the deflection profile due to a moving load is well established¹⁻³, there has been little progress experimentally and early studies^{4,5} were subject to error because of the type of transducer used. Recently there has been a renewed interest in the problem which has led to several theoretical papers⁶⁻⁹ and to the collection of a small quantity of high-quality data^{10,11}. We report here some preliminary results from a new and complete set of experiments done on Antarctic sea ice using strain gauges to measure directly the strain induced by the vehicle. The results show excellent agreement with theory in all respects.

The project was staged from a small camp on the sea ice near Tent Island in McMurdo Sound. This area was chosen because the ice was particularly smooth at its upper surface, enabling the vehicle to be driven at high speed, and it was consistently 1.6 m thick over a large area. Uniform thickness is important because the rigidity of the ice is proportional to the cube of thickness. The vehicles used in the experiments were an extended cab pickup truck of weight 2,100 kg and an LC-130 Hercules aircraft weighing some 50,000 kg. Strain gauges designed at the University of Cambridge and at the Physics and Engineering Laboratory were used to record induced surface strain, and were

found to agree to within better than 5%. Vehicle experiments took place along a straight 6 km-long ice road, and ample distance was allowed for the deflection field due to the vehicle to become steady. All distances were found by laser survey.

Thickness, h , was measured directly by drilling. Core drilling was used to determine salinity, temperature (and hence brine volume) and crystal fabric profiles; a thermistor probe frozen into the ice monitored temperature. These data enabled the effective modulus, $E(z)$, and hence the effective integrated flexural rigidity, D , for the ice to be determined. It is necessary to use effective modulus because our experiments are in a high temperature-low strain rate regime. The true Young's modulus, which governs deformation at very high strain rates and very low temperatures, is not directly relevant¹². Effective modulus is a function of rate and temperature, but for non-saline ice there are data relating the effective modulus to porosity, strain rate and temperature. We may therefore calculate the effective modulus profile for non-saline ice with an identical temperature profile at the loading rates of our experiment. From this we find the effective modulus profile for our sea ice, and then integrate to obtain the neutral axis ($z_0 = 0.79$ m) and the effective flexural rigidity

$$D = \frac{1}{1-\mu^2} \int_{-z_0}^{h-z_0} z^2 E(z) dz \quad (1)$$

where μ is effective Poisson's ratio. D was found to be 1.8 ± 0.3 GNm, assuming $\mu = 0.3$, as is usual for quasi-elastic sea-ice problems.

Strain gauges, set in 45° rosettes to enable principal strains and their directions to be found, were deployed alongside the road and respectively at 30, 100, 200, 400 and 800 m from the road's centre line. In each experiment the vehicle was driven along the road at 10, 20, 30, 35, 38, 40, 45, 50, 55 and 60 miles per hour (m.p.h.) as measured by the truck's speedometer, and the induced strain measured at the various strain gauge sites. Accurate timing over a known distance was done to establish the vehicle's true average speed. For experiments involving strain rosettes located further than 30 m from the road, the slower speeds were omitted as induced strain at the rosette was negligible. In the aircraft experiments, the altitude of the LC-130

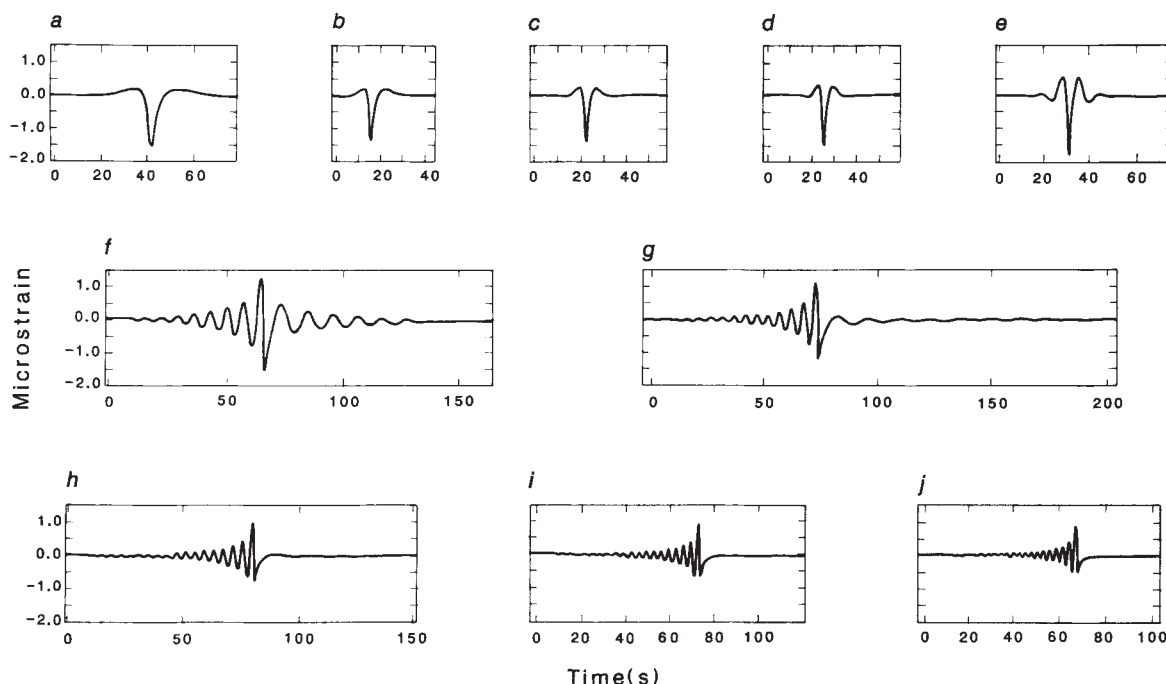


Fig. 1 Strain gauge records for different speeds recorded on an instrument alongside and parallel to the ice road. Strains are in microstrain (10^{-6}), and speeds in ms^{-1} are as follows: a, 4.5; b, 8.9; c, 13.8; d, 15.7; e, 17.5; f, 18.4; g, 20.8; h, 23.2; i, 25.8; j, 28.2. Time zero is arbitrary.

At high speeds the vehicle lags the maximum negative strain by a small amount.

was held at 50 feet (15.2 m) and three speeds were flown at our request, namely 100, 130 and 160 knots. Strain rosettes were located at 0 and 800 m.

There are two regimes of interest: a slow speed regime, where speed u is less than some critical velocity $c_{\min}$ and no waves are produced; and a high speed regime, where $u > c_{\min}$ and two sets of waves are generated, one of short wavelength leading the vehicle and one of longer wavelength trailing the vehicle. Other critical velocities can exist if, for example, the waves touch bottom or if the vehicle loading on the ice is oscillatory. The leading wave is controlled predominantly by ice flexure and the trailing wave by the fluid beneath the ice, although as one approaches $c_{\min}$ the distinction is less clear and the two wavelengths approach the same limiting value. When $u = c_{\min}$ no energy escapes from beneath the vehicle and resonance occurs.

When $u < c_{\min}$ the behaviour is quasi-static and we may draw on the early literature for stationary loads on thin elastic ice sheets^{13,14}. Directly beneath the load, the surface strain magnitude is

$$\frac{3fg(1+\mu)l \operatorname{kei}'(R/l)}{ER\pi h^2} \quad (2)$$

where f is the load which acts over a radius R , g is acceleration due to gravity, and $l = D/\rho'g$ is the characteristic length where ρ' is the density of the sea water foundation. The wavelength, λ , of the strain field is given approximately by the solution of

$$\ker(\lambda/2l) = \frac{2l}{\lambda}(1-\mu) \operatorname{kei}'(\lambda/2l) \quad (3)$$

which gives $\lambda = 4.2l$. It is found that the fully static solution agrees well with observations for $u < c_{\min}$.

When $u > c_{\min}$, a dispersion equation is written down which again assumes that the sea ice behaves as a thin elastic plate on a fluid foundation¹⁵. For algebraic simplicity we present here the deep water equation in terms of the phase velocity c ,

$$c^2 = \frac{Dk^3 + \rho'g/k}{\rho h k + \rho'} \quad (4)$$

where $k = 2\pi/\lambda$ is the wavenumber and ρ is the ice density. In the viscoelastic case D would be complex. Expression (4) has a minimum, $c_{\min}^2 = 4Dk^3/(2\rho h k + \rho')$, below which no steady wave train can propagate; $c_{\min}$ is the critical velocity referred to above. Equation (4), which has two sensible roots for k and hence λ for each c , enables wavelength to be plotted as a function of vehicle speed, u .

In Fig. 1 we have plotted calibrated data from one of the experiments, recorded at sub-critical and super-critical speeds. Note the sudden change in the form of the strain signal as the critical velocity is exceeded; no waves when $u < c_{\min}$, and waves before and after the load when $u > c_{\min}$. Also note the gradual increase in strain amplitude from its static value up to that at critical velocity, and then its decrease beyond. This effect may be seen more clearly in Fig. 2, which shows a distinct resonant peak at $18 \pm 0.5 \text{ ms}^{-1}$. The variation in Fig. 2 between experiments derives from increased compliance in the ice due to higher temperatures and fits thermistor probe data; the position of $c_{\min}$ is unaffected by this variation. Using $c_{\min}$ we have an alternative way of finding the effective flexural rigidity D by means of the elastic thin plate equation. When this is done we find $D = 1.6 \pm 0.3 \text{ GNm}$, in good agreement with our earlier estimate of $1.8 \pm 0.3 \text{ GNm}$ based on temperature and salinity measurements. The peak of the curve in Fig. 2 of ≈ 2.8 times the static value is higher than the values of 1.45 and 2.25 determined previously on snow-covered sea ice and lake ice¹⁰. The lower value for sea ice reported in earlier work is most likely to be due to the damping effect of the $\approx 0.5 \text{ m}$ layer of snow.

Figure 3 shows three strain gauge records collected during the aircraft experiments. The records have been low-pass filtered,

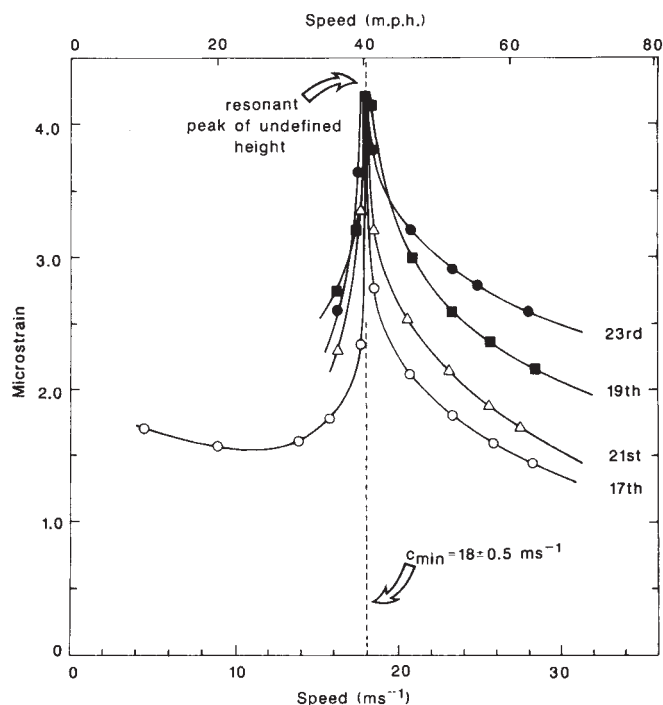


Fig. 2 Strain excursion between maximum tensile and compressive strains as a function of vehicle speed. Curves are marked with the date of the experiment.

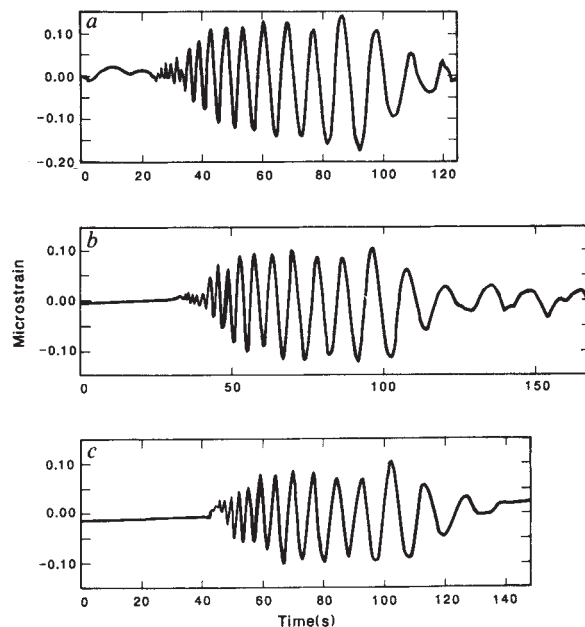


Fig. 3 Strain gauge records 800 m from the ice road and aligned at 90° to its centreline, for an LC-130 Hercules aircraft flying at 50 ft at: a, 100; b, 130; and c, 160 knots above the road.

as the raw data were affected slightly by a small, periodic spike from the aircraft's radar beacon; the beacon could not be switched off for safety reasons. The figure shows the variation in amplitude and wavelength of dispersive wave packets generated by the aircraft. The shorter, flexurally controlled waves lead the longer, gravity-controlled waves with a factor 6 increase in wavelength. In Fig. 3a-c the amplitude of the waves decreases inversely with speed, with roughly 30% decrease between 52 and 83 ms^{-1} .

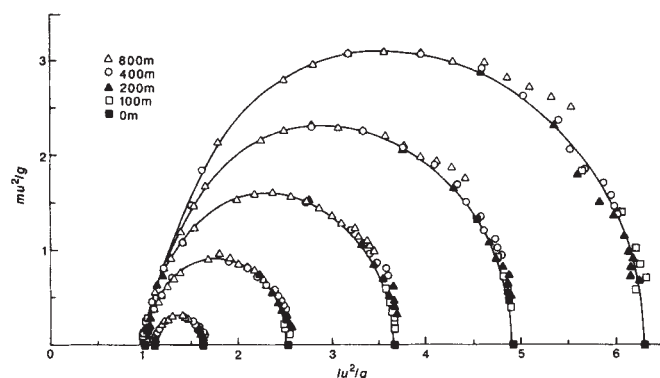


Fig. 4 Wavenumber curves derived from strain gauge records at 0, 100, 200, 400 and 800 m from the road for vehicle speeds of 18.5 (inner curve), 20.7, 23.3, 25.7 and 28.0 m s⁻¹ (outer curve).

Equation (4) may be generalized to two dimensions as follows

$$c^2 = \cos^2 \beta = \frac{ak^4 + 1}{bk^2 + k} \quad (5)$$

where $a = Dg^3/u^8\rho'$, $b = gh\rho/u^2\rho'$, and $(k \cos \beta, k \sin \beta)$ is the wavenumber vector. The expression is normalized such that $ku^2/g \rightarrow k$ and $c/u \rightarrow c$. In Fig. 4 we have plotted these wavenumber curves for speeds corresponding to the experiments, assuming $D = 1.6 \text{ GNm}$. Also plotted are the set of data points for the appropriate speed that are found from the wavecrests seen by strain gauges located at all of the sites except that at 30 m. 30 m points are excluded (despite their lying perfectly on the curve) to avoid clutter. There is excellent agreement between data and theory. The slight lifting of some of the data points away from the curve may be due to flexural hysteresis in the ice, as these points derive from wavecrests well ahead of the vehicle where such effects have had time to act.

In conclusion, we note that a single strain gauge frozen into the McMurdo Sound sea ice runway, or indeed any ice road or runway, could be used to monitor routinely the effective flexural rigidity D for the sea ice via a thin plate dispersion equation. As D is an integrated quantity, it contains information about the state of the ice right through its thickness; from skeletal layer to its upper surface. Furthermore, when the ice is cold and hard, resonance will be nearly undamped. As the ice warms, and presumably becomes less safe, so the resonance will be more heavily damped. This leads to the idea of adopting D and the quality 'Q' of the resonant peak as indicators of ice condition. This is equivalent in a sense to using complex compliance, but includes a thickness factor.

In principle, therefore, a strain gauge offers an effective and inexpensive sensor for determining the condition of the ice at the strain rates produced by landing aircraft and by vehicles operating on ice roads. Note also that as we do not yet have a reliable model for sea ice creeping under a static load, such as parked aircraft, the strain gauge could be further used to provide direct readings of creep deflection in real time.

We are indebted to the NZ Ross Dependency Research Committee, DSIR, the University of Cambridge, Newnham College, the Royal Society of London, and the Trans-Antarctic Association for supporting this work, and to the US Navy who operate the National Science Foundation's LC-130 Hercules aircraft. We also thank Arnold Heine, our field technician.

Received 22 December 1987; accepted 23 March 1988.

- Assur, A. *SIPRE Rep.* 36 (1956).
- Assur, A. *Proc. 1st Inter. Conf. Mech. Soil-Vehicle Systems* 913-923 (Torino, Italy, 1961).
- Nevel, D. *CRREL Res. Rep.* 261, 1-13 (1970).
- Eyre, D. J. *Glaciol.* 19, 555-570 (1977).
- Beltaos, S. *Can. J. Civ. Eng.* 8, 1-8 (1980).

- Kerr, A. D. *Cold Reg. Sci. Tech.* 6, 267-274 (1983).
- Davys, J. W., Hosking, R. J. & Sneyd, A. D. *J. Fluid Mech.* 158, 269-287 (1986).
- Schulkes, R. M. S. M. & Sneyd, A. D. *J. Fluid Mech.* (in the press).
- Hosking, R. J., Sneyd, A. D. & Waugh, D. W. *J. Fluid Mech.* (in the press).
- Squire, V. A., Robinson, W. H., Haskell, T. G. & Moore, S. C. *Cold Reg. Sci. Tech.* 11, 123-139 (1985).
- Takizawa, T. *Cold Reg. Sci. Tech.* 11, 171-180 (1985).
- Mellor, M. *CRREL Monogr.* 83-1, 105 pp (1983).
- Hertz, H. *Ann. Phys. Chem.* 22, 449-455 (1884).
- Wyman, M. *Can. J. Res.* 28, 293-302 (1950).
- Greenhill, A. G. *Am. J. Math.* 9, 62-112 (1887).
- Lighthill, J. *Waves in Fluids* (Cambridge University Press, London, 1978).

The underside of Arctic sea ice imaged by sidescan sonar

Peter Wadhams

Scott Polar Research Institute, University of Cambridge,
Lensfield Road, Cambridge CB2 1ER, UK

Despite extensive submarine operations in the Arctic, little is yet known about the three-dimensional morphology of the ice underside. Several studies have been done using upward-looking sonar¹⁻⁷, which yields a linear profile of ice draft from which the distributions of features such as pressure ridges and leads may be derived. Such a profile gives no information about the spatial morphology or appearance of the ice bottom surface. In contrast, the upper ice surface has been well documented from airborne and satellite-borne sensors such as synthetic aperture radar, passive microwave, infrared and visual imaging. Here I report the first extensive set of imagery of the underside of sea ice in the Arctic Ocean. Sidescan sonar images show a distinct difference between the appearance of first-year ice and of multi-year ice. First-year ice has a smooth bottom except for a network of cracks, while multi-year ice is covered with an array of blisters or bulges. The clear discrimination between ice types enables sidescan sonar to be used as a means of validating airborne passive microwave data.

A single previous instance of under-ice imaging was an attempt to obtain sidescan sonar imagery from HMS *Sovereign* in 1976^{8,9}, when a limited dataset was obtained from one channel of a sidescan transducer over a distance of 40 km before the transducer flooded. In May 1987 HMS *Superb* carried out an extensive ice profiling experiment in the Arctic Ocean, during which she obtained >6,000 km of upward-looking sonar profile and 3,000 km of sidescan sonar imagery. The experiment was co-ordinated by the author, who sailed aboard the submarine. After surfacing at the North Pole on 18 May *Superb* undertook four 24-hour legs of ice profiling: down the 55° W meridian to 85° N; eastwards to 84° N 10° W; down into Fram Strait at 79° N 0° W; and through the Greenland Sea to 75° N 10° W. Each leg was overflown on the day of completion by two collaborating remote sensing aircraft: a Cessna Conquest of Intera Technologies Ltd, Calgary, equipped with synthetic aperture radar (SAR); and a Lockheed P-3 of NASA equipped with passive microwave radiometers at three frequencies (18, 21 and 37 GHz), a scanning microwave radiometer, a laser profilometer, an infrared radiometer and cameras. The purpose of the experiment was to compare the under-ice record obtained by the submarine with the data obtained by the airborne sensors along identical tracks at near-identical times, as a way of achieving a properly controlled validation of these sensors (particularly passive microwave and SAR). In this respect the data obtained by the sidescan sonar were of crucial importance.

The submarine was equipped for the experiment with an EDO Western model 602 sidescan sonar towfish, mounted on the upper casing aft of the fin, with a tubular steel cage over it to protect against ice damage. The towfish fed an EDO 706 sidescan mapping system operating at 100 kHz. This system has a slant

Fig. 1 A section of sidescan sonar profile from the Arctic Ocean showing two multi-year floes. Transducer depth 75 m. A and B are features discussed in the text.

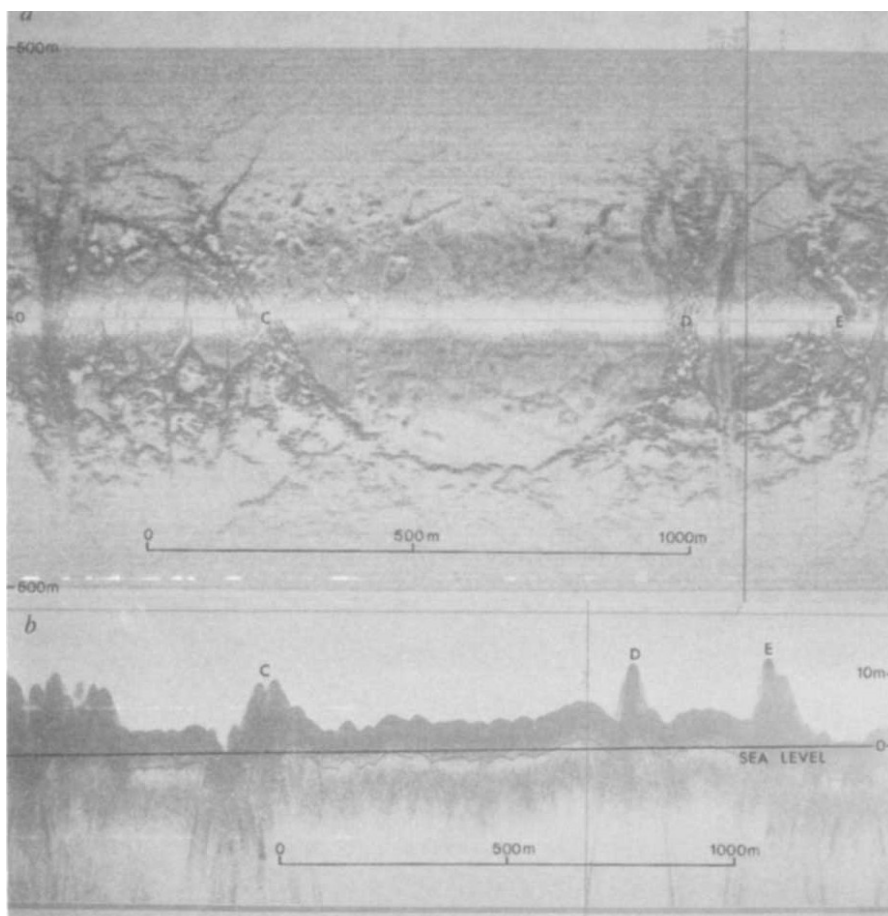
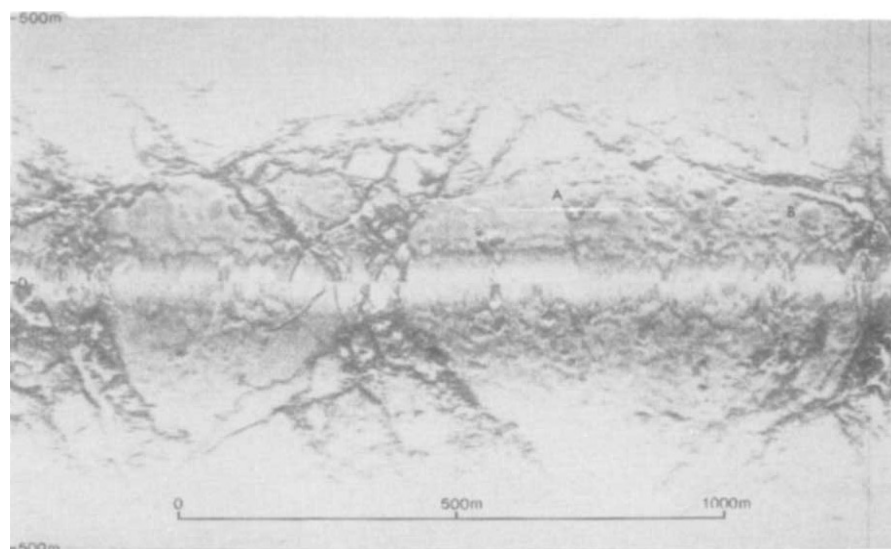


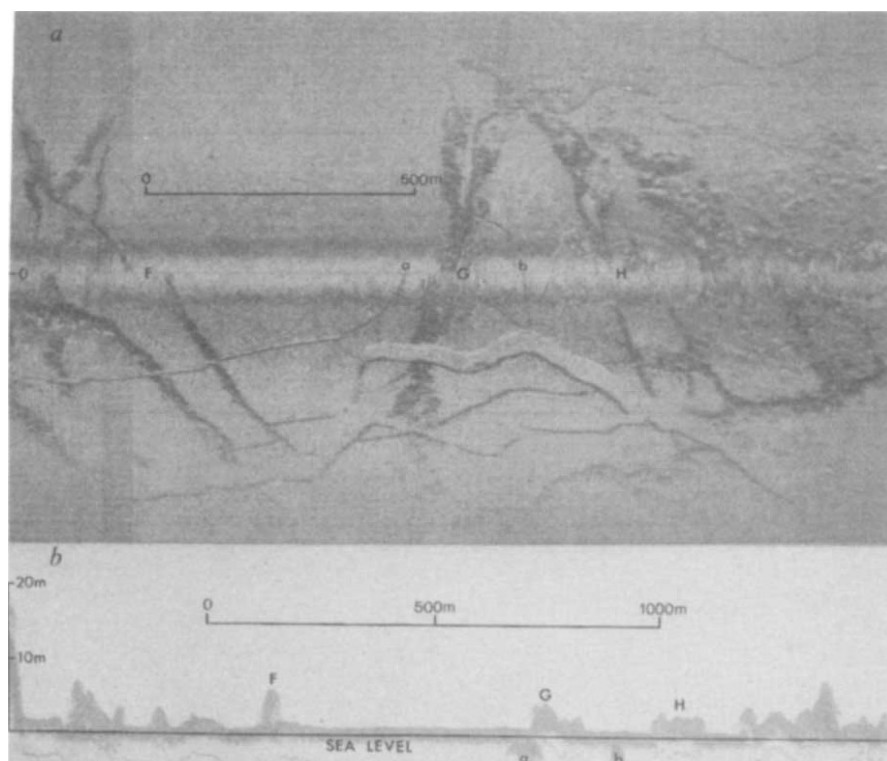
Fig. 2 *a*, Multi-year floe in a sidescan profile from north of Greenland. *b*, The corresponding upward sonar profile of ice draft. Ridges C, D and E are identifiable on the sidescan image.

range correction facility and a variable paper speed keyed to ship speed, so that at speeds of <8 kt ($1 \text{ kt} = 1.852 \text{ km h}^{-1}$) it is possible to obtain a geometrically correct map of the ice underside along the submarine's track out to a lateral distance of up to 600 m. It was found that the practical range to which good-quality data were obtained was 300–350 m, giving an overall swath width of 600–700 m. Data were recorded on paper chart and analogue magnetic tape. The submarine was also equipped with a 48-kHz upward-looking echo sounder and an upward-looking TV system.

Here I present some sidescan sonar images that demonstrate hitherto unsuspected characteristics of the ice's underside. The most important discovery of the cruise was that undeformed first-year ice and undeformed multi-year ice have quite different qualities of roughness and morphology on their underside, and are thus easy to discriminate.

Figure 1 shows two typical multi-year floes, profiled at $86^{\circ} 36' \text{ N}$, $54^{\circ} 30' \text{ W}$, with the submarine running at a transducer depth of 75 m. The floes are of dimensions $880 \text{ m} \times 620 \text{ m}$ and $430 \text{ m} \times 500 \text{ m}$, and are bounded and separated by the keels of

Fig. 3 *a*, A first-year floe near the North Pole, with a refrozen lead that has opened from one of the cracks in it. *b*, The corresponding upward sonar profile. Ridges F, G and H are common. The cracks *a*, *b* are visible as hyperbolae in the upward sonar profile.



pressure ridges. Dark areas on the image represent strong sound returns, generated by rough surfaces or by planes inclined towards the fan-shaped beam emitted by the sidescan (mainly the nearside flanks of pressure-ridge keels). White areas are shadow zones, typically on the far sides of keels. Thus the pattern of light and shadow in the image is the negative of that of a surface illuminated by a strip light running along the centreline of the image at an elevation of 75 m.

The dominant feature of these floes is a characteristic random pattern of blisters or bulges that protrude from the floe bottom in a gentle but unmistakable way. The protuberances are smooth-surfaced, with diameters ranging from 14 m for the smallest to 80 m for the largest on that floe (feature A). They are often almost circular, with feature B as an especially clear case. The whole floe resembles the surface of the Moon with the craters turned inside out. The disposition of blisters over the surface appears random.

Figure 2*a* shows a similar multi-year floe surrounded by rough deformed ice. It was situated at 85° 01' N, 38° 20' W, in a region just north of Greenland which is known for its heavy pressure ridging³. Again, the surface of the floe is covered with random bulges with a wide range of diameters but an apparently more muted relief than the floe in Fig. 1. The corresponding upward sonar profile (Fig. 2*b*), giving ice drafts along the centreline of the sidescan image, is registered by the locations of ridges C, D and E, which delineate the edges of the main floe and of the smaller floe beside it. It can be seen that the main floe has a mean draft of about 5 m; the blisters have wavelengths in the range 28–63 m, and peak-to-trough reliefs of 0.5–1.2 m.

In complete contrast, Fig. 3 shows a first-year floe situated near the North Pole at 89° 29' N, 55° W. It is criss-crossed by pressure ridges; the apparent discontinuity of ridge F as it crosses the centreline is probably caused by a slight list to the submarine, which would result in the slant range correction software making an error in the mutual registration of the port and starboard channels. The floe bottom is extremely smooth, except for a number of narrow line cracks traversing the floe. One of these has widened to form a refrozen lead of width 35 m, lying away

from the centreline of the image. The corresponding upward sonar profile (Fig. 3*b*) shows that the mean draft of the floe is 1.5 m, and that the ridges themselves are quite shallow, with reliefs of only 3.5–5.5 m. The featureless sidescan return corresponds to the smooth upward sonar record. Two of the narrow cracks traversing the floe (*a*, *b*) show up in the upward sonar record as hyperbolae caused by the effect of sonar beamwidth.

Finally Fig. 4 is another first-year floe, also near the North Pole at 89° 42' N, 58° W. Once again the floe bottom is smooth and featureless except for the pattern of narrow cracks, none of which has opened up to form a lead, however. One of the cracks has a very clear white line on the nearside with black on the far side, indicating that it has definite width and depth. The width must exceed the resolution limit of the sidescan, about 3 m. The corresponding upward sonar profile (Fig. 4*b*) again shows a floe of mean draft 1.5 m, with a smooth bottom except for a number of hyperbolic echoes from cracks.

What is the origin of these differing topographies? The smooth first-year ice bottom surface is as expected, the only surprise being the pattern of narrow cracks that is always present. This pattern is not normally apparent on the top surface of first-year floes, probably because it is always obscured by the snow cover. Each crack can be seen as a potential lead, and the pattern may be caused by mechanical stresses exerted from the thicker pressure ridges and surrounding rough ice, or it may be the result of thermal stresses set up in the first-year floe itself by the vertical temperature gradient within it.

The rough surface of multi-year floes is more difficult to explain. The two possible hypotheses are that the bulges represent worn-down pressure ridges, that is, they are the remnants of deformation of the floe occurring in past years, or that they are the product of the melt-freeze cycle that the floe has undergone each summer. The first hypothesis appears unlikely, because there is no apparent lineation to the bulges as might be expected if they were the relics of old ridges. It is known that multi-year floes have a rough upper-surface topography of hummocks and depressions. The conventional explanation of their origin¹⁰ is that the depressions represent sites at which

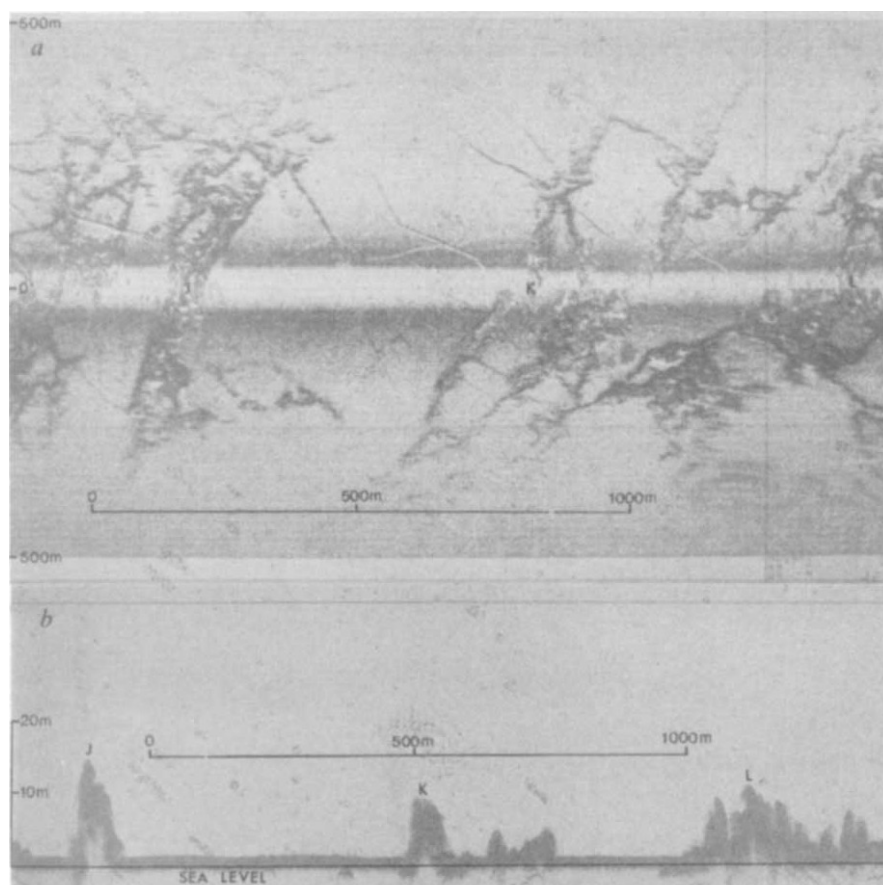


Fig. 4 *a*, A first-year floe near the North Pole. *b*, The corresponding upward sonar profile.

melt ponds were located during the summer. During a floe's first summer of melt the surface meltwater gathers into ponds which form wherever the freeboard is very slightly lower than average. These initial very small variations in freeboard then increase as the meltwater pools absorb solar radiation and grow deeper. Bottom melting also proceeds more rapidly under the meltwater pools so that each depression in the top surface is mirrored by a cavity in the bottom surface. Meltwater may drain through to the underside of the floe where it forms an under-ice melt pond¹¹, a layer of fresh water 0.5–1 m thick between the ice bottom and the normal sea water. Such a pond will then refreeze in autumn to add a layer to the bottom of the ice. There is some question¹² as to whether drainage commonly occurs through a continuous ice sheet or mainly around the edges of floes and through cracks. The topography variations generated by this first summer of melt are then perpetuated and amplified in subsequent summers, when melt ponds will form in the same locations on the floe.

There has been very little direct measurement of the bottom shape of multi-year floes to test these ideas. Cox and Weeks¹⁰ drilled a series of cores across a hummock–depression–hummock sequence on a multi-year floe, finding a mirror topography on the underside, with a relative relief of 1 m in a floe of 2–3 m draft, the hummock having a wavelength of 12 m. This appears to be compatible with the features which we see

on the sidescan image. However, a rolling topography of hummocks and depressions caused by the melt-freeze cycle would be expected to yield a gaussian surface with 'pits' as well as 'bulges'. The appearance of the images is that of a flat surface on which only positive bulges have been superposed. The cause of this type of topography remains to be understood. A tentative hypothesis is that the under-ice freshwater layers of late summer may refreeze discontinuously in the form of lenses which are the origin of the bulges.

Much work remains to be done with this very large data set. The sidescan reveals details of ridge structure, of floe shapes near the ice edge, and of frazil ice bands outside the ice edge. But the major feature seen in the records is the clear first-year/multi-year ice distinction, such that submarine sidescan sonar can indeed be used as a validating procedure for passive microwave and SAR. This will be of great importance in the development of valid algorithms for retrieving the fraction of multi-year ice from the data now being received from the SSM/I (special sensor microwave imager) on the DMSP satellite launched in June 1987.

I am grateful to the Ministry of Defence (Navy) for release of data; to Flag Officer Submarines for the opportunity to take part in the voyage; to the Commanding Officer and crew of HMS *Superb*; and for financial support to the Office of Naval Research through a grant to Polar Oceans Associates.

Received 22 December 1987; accepted 25 February 1988.

- Williams, E., Swithinbank, C. W. M. & Robin, G. de Q. *J. Glaciol.* **15**, 349–362 (1975).
- Wadhams, P. & Horne, R. J. *J. Glaciol.* **25**, 401–424 (1980).
- Wadhams, P. *Phil. Trans. R. Soc. A* **302**, 45–85 (1981).
- Wadhams, P. *Nature* **305**, 108–111 (1983).
- McLaren, A. S., Wadhams, P. & Weintraub, R. *Arctic* **37**, 110–120 (1984).

- Wadhams, P., McLaren, A. S. & Weintraub, R. *J. geophys. Res.* **90**, 1069–1077 (1985).
- Wadhams, P. & Davy, T. *J. geophys. Res.* **91**, 10697–10708 (1986).
- Wadhams, P. *Can. J. Remote Sensing* **4**, 161–173 (1978).
- Wadhams, P. *J. Soc. Underwat. Technol.* **9**, 1–12 (1983).
- Cox, G. F. N. & Weeks, W. F. *J. Glaciol.* **67**, 109–120 (1974).
- Martin, S. & Kauffman, P. *J. Fluid Mech.* **64**, 507–527 (1974).
- Weeks, W. F. & Ackley, S. F. in *The Geophysics of Sea Ice* (ed. Untersteiner, N.) 9–164 (Plenum, New York, 1986).

Metal remobilization at a spreading centre studied using lead isotopes

Bernard Dupré*, Gérard Blanc†, Jacques Boulègue† & Claude J. Allègre*

* Laboratoire de Géochimie et de Cosmochimie, IPG tour 14-24, 4 place Jussieu, 75252 Paris Cedex 05, France

† Laboratoire de Géochimie et de Métallogénie, Université Pierre et Marie Curie, Tour 26, 4 place Jussieu, 75252 Paris Cedex 05, France

The metalliferous sediment deposits of the Atlantis II Deep in the Red Sea are an example of submarine hydrothermal deposits at a spreading centre. The brine salinities (~270‰ for the Atlantis II Lower Brine) originate from leaching of the Miocene evaporites. The mantle is considered as a possible source of metal by analogy with East Pacific Rise deposits. During the 'Hydrotherm' cruise (May, 1985) on *RV Marion Dufresne*, two cores, numbered 683 and 684 were collected from the SW and W basins respectively of the Atlantis II Deep (Fig. 1). These two cores reached the basaltic basement and therefore allowed full and representative sampling of the metalliferous sediment series, as well as of the underlying basalt. Our lead isotope results suggest that the metals in the sediments are not derived from a basaltic source. We show that lead in the metalliferous sediments was derived both from the Recent biogenic-terrigenous sediments and another unknown sedimentary source. The variations in the lead isotopic composition of Atlantis II Deep sediments suggest a discontinuous brine supply.

Previous results on cores^{1,2} allowed the establishment of lithological zonations complementary to those defined by Bäcker and Richter³. Figure 2 shows the simplified lithology of these cores and the locations of the sediment samples. We also analysed the basalt and glass samples described at the bottom of these cores.

Core 684 (W basin). Carbonates (zone E) at the bottom of the sedimentary series (Fig. 2) were derived from biological and detrital sedimentation before the first supply of a metal-bearing fluid. Radiocarbon dating⁴ indicates that this biogenic-terrigenous sedimentation began about 28,000 yr ago. The E-zone average sedimentation rates were about 25 cm per thousand years. These rates are lower than the overall hydrothermal sedimentation rates (100 cm per thousand years), but higher than the overall bio-detrital sedimentation rates in the Red Sea, which range from 2 to 15 cm per thousand years^{5,6}. The occurrence of goethite-rich horizons at the upper portions of the carbonate unit is evidence of the beginning of supply of a metal-bearing fluid to the upper part of zone E.

Core 683 (SW basin). This core contains three main zones which, from the bottom to the top, are: (1) zone I (16.20–11.75 m). The most abundant minerals are iron oxides and sulphates deposited under oxidizing conditions that occurred in contact with the Red Sea Deep Water. At a later stage a thin brine layer floated on top of the layers. (2) Zone ST (11.75–10.80 m). Sulphates are the main minerals found in this zone. The conditions during mineral deposition became more reducing as the brine thickness increased over the sediment. (3) Zone S (10.80 m from the top). Sulphide minerals are very abundant and sulphates occur sporadically. The oxides and hydroxides present are hematite (between 7.85 and 8.85 m) and goethite (between 8.85 and 10.00 m). Their deposition could be related to a quiescence of the influx of metal-bearing fluid into the Atlantis II Deep².

Additional sampling. To define the lead isotopic composition ranges of the Red Sea basalt, we analysed several basalts from Red Sea Ridge (Table 1). Also, two Recent sediments (<25,000 years), SIPAN 2 and SIPAN 38, collected in the northern part of the Atlantis II Deep (latitude 21°48' N, longitude 37°56' E) during the MD29 cruise of the *Marion Dufresne* were analysed. Additional sampling included four samples from the

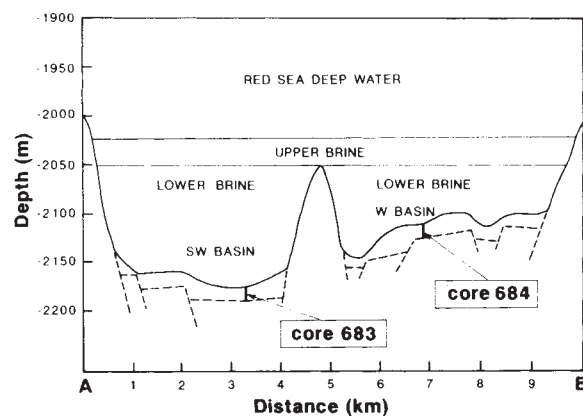


Fig. 1 General profile of (SW and W basins), showing cores 683 and 684.

Miocene evaporite series (DSDP leg 23), analysed to evaluate the evaporites' sequence as a potential source of lead for the metalliferous Atlantis II Deep sediments. They are a black shale (gravelly sand) (SED 227-40-2, 55-57) (21°19.86' N; 38°07.97' E); one evaporite with anhydrite, dolomite and magnesite (SED 225-27-2, 33-35) (21°18.58' N; 38°15.11' E); one evaporite with anhydrite and halite (SED 225-26-2, 53-55), and one sample of pure halite (SED 225-29-3, 113-115).

The results of this lead isotopic study are given in Table 1, using the chemistry described by Manhès *et al.*⁷. Blank and standard analyses are given in the Table 1 legend. The isotopic compositions of the basalts were determined after leaching with 6 M HCl. The $^{206}\text{Pb}/^{204}\text{Pb}$ ratio for the Red Sea Rift basalts ranges from 19.163 to 18.263, whereas the $^{206}\text{Pb}/^{204}\text{Pb}$ ratios of Atlantis II basalts range from 18.263 to 18.354. The lead isotopic compositions of the glass and altered basalt sample show no significant differences. Thus, the $^{206}\text{Pb}/^{204}\text{Pb}$ average ratio of the Atlantis II Deep basalts is ~18.3, and the $^{208}\text{Pb}/^{204}\text{Pb}$ ratio of 38.1 is similar to the Pb isotopic composition found in Indian Ocean tholeiites⁸⁻¹⁰. The average $^{206}\text{Pb}/^{204}\text{Pb}$ ratio of the Recent-sediment samples is about 18.85, which is very different from the 18.3 value obtained for the Atlantis II area basalts. But the Recent sediment $^{206}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios are in good agreement with the pelagic sediment lead isotopic values from site 227 given by Delevaux and Doe¹¹. Four metalliferous samples from core 683 gave $^{206}\text{Pb}/^{204}\text{Pb}$ ratios of very limited variability (18.702–18.742). The highest value occurs in a sediment layer formed during a quiescent period in the supply of metal-bearing fluid². The core 684 sediment samples show more scattering of lead isotopic values than those from core 683 (for example, $^{206}\text{Pb}/^{204}\text{Pb}$ ratios range from 18.696 to 18.823). The two carbonates collected at the bottom of core 684 are the most radiogenic samples and show large variations of the $^{208}\text{Pb}/^{204}\text{Pb}$. In sample 684-7, the influence of the brine leads to a lower value of $^{208}\text{Pb}/^{204}\text{Pb}$ ratio. The isotopic value for the carbonate sediment 684-8 is close to that measured in the Recent pelagic sediments from the vicinity of the Atlantis II Deep. These sediments therefore apparently have the same source of lead. The $^{206}\text{Pb}/^{204}\text{Pb}$ ratio correlates with depth (Fig. 3). The metalliferous sediments are less radiogenic than the carbonate sediments sampled at the bottom. All metalliferous sediments coming from 683 and 684 are on a line in both lead diagrams: $^{208}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ (Fig. 4) and $^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ (not shown). The Miocene evaporite having anhydrites, dolomite and magnesite gave a value of 18.705 for the $^{206}\text{Pb}/^{204}\text{Pb}$ ratio and 38.175 for the $^{208}\text{Pb}/^{204}\text{Pb}$ ratio. These Pb isotopic ratios are not too different from the dark siltstone values¹¹. In contrast, the black shale has a very radiogenic lead composition. Lead concentrations of the pure halite sample and halite with anhydrite sample are very low (<1 ng g⁻¹) and the lead isotopic compositions were not determined.

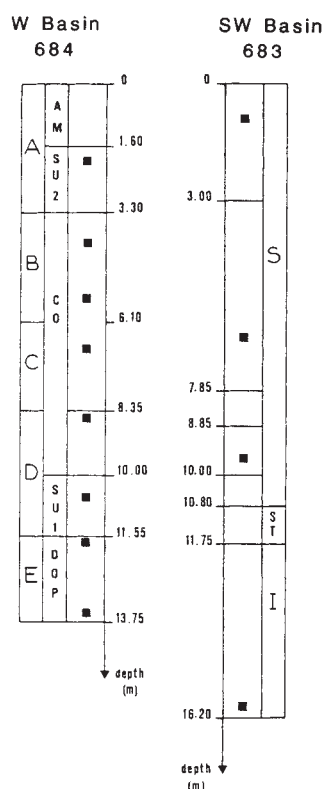


Fig. 2 Sample location in cores 683 and 684. Zonation for core 684(A-E)¹ is compared with zonation from B  cker and Richter³. In core 683 these zonations do not apply and sedimentological zonation by Blanc *et al.*¹ is given (see text).

One result of our study is that the reservoir rocks (basalts, recent sediments and evaporites series) that can participate in the genesis of the metalliferous sediments from Atlantis II Deep must have very different lead isotopic compositions. Another significant result is that all the metalliferous sediment samples

show a linear relationship in the $^{206}\text{Pb}/^{204}\text{Pb}$ versus $^{208}\text{Pb}/^{204}\text{Pb}$ diagram (Fig. 4) and in the $^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ (not shown). Therefore, lead found in the metalliferous sediments can be interpreted as resulting from the mixing of two sources. Basalts sampled in the Atlantis II Deep deviate substantially from the linear relationship. Therefore, most of the lead in the metalliferous sediments is not derived from basalts. The most radiogenic end-member can be identified. It corresponds to the carbonate sediments found at the basis of 684 core or to the Holocene sediments found north of Atlantis II. The second end-member has a less radiogenic source of lead ($^{206}\text{Pb}/^{204}\text{Pb} = 18.7$; $^{207}\text{Pb}/^{204}\text{Pb} = 15.52$; $^{208}\text{Pb}/^{204}\text{Pb} = 38.2$). The samples of the evaporite sequence that we have analysed are not on the mixing line; the concentrations of lead in anhydrite and halite are also very low (<1 ng), which excludes them as a possible source. We analysed only two samples in the evaporite sequence, however, and these have a very different isotopic composition; we propose that the second end-member is located in the sedimentary column, either in the Miocene evaporite or in the Pleistocene sediments in the Red Sea. One argument in favour of this solution is the result obtained for the dark siltstone¹¹, whose isotopic value plots near the value for the second end-member. The metal-to-iron weight ratio in the metalliferous sediments of Atlantis II are¹²: $\text{Zn}/\text{Fe} = 1.2 \times 10^{-1}$, $\text{Cu}/\text{Fe} = 4.5 \times 10^{-1}$ and $\text{Pb}/\text{Fe} = 3.5 \times 10^{-3}$, well above the corresponding values for basalts¹³ (2×10^{-3} , 1.5×10^{-3} , 10^{-5}) or for the hydrothermal sediments^{14,15} ($\text{Zn}/\text{Fe} = 3 \times 10^{-3}$, $\text{Cu}/\text{Fe} = 4 \times 10^{-3}$). They are close to the ratio found in biogenic sediments¹⁵: $\text{Zn}/\text{Fe} = 8 \times 10^{-2}$, $\text{Cu}/\text{Fe} = 5 \times 10^{-2}$. This suggests that in Atlantis II Deep, metals other than lead do not have a basaltic source, and probably result from leaching of the Red Sea sediments.

The mixing between two end-members for lead can be the result of three distinct processes acting during sedimentation in the Atlantis II Deep. (1) The recharge fluid, for example sea water, leaches sediments in its downward path. Lead and other

Table 1 Lead isotopic compositions of basalts, Atlantis II Deep metalliferous sediments, Recent sediments and Miocene evaporites

Sample	Type	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$	Concentration (p.p.m.)	Depth (m)
Basalts						
Mahabis Deep	KS 11-A	18.719	15.531	38.514	0.2	
	DR-07-E	18.735	15.559	38.690	0.4	
	KS-04-A	18.779	15.569	38.690	0.2	
18° N	MR 80 53-2	19.116	15.590	38.857		
	MR 80 59-3	18.787	15.555	38.575		
	MR 80 68-2	19.141	15.577	38.876		
	MR 80 71-2	19.163	15.588	38.895		
Atlantis II Deep	684 ogive(glass)	18.263	15.496	38.024	0.1	
	684 ogive	18.284	15.505	38.098	0.1	
	399VII 53(glass)	18.354	15.515	38.143	0.1	
Metalliferous sediments						
684-1	Metal. sed.	18.744	15.545	38.319	21.4	2.36
684-2	Metal. sed.	18.696	15.531	38.251	1.6	4.15
684-2	Metal. sed.	18.699	15.541	38.279	1.7	4.15
684-3	Metal. sed.	18.756	15.573	38.396	10.4	5.5
684-4	Metal. sed.	18.735	15.550	38.330	8.9	6.84
684-5	Metal. sed.	18.768	15.580	38.450	487.0	8.56
684-6	Metal. sed.	18.764	15.560	38.378	32.6	10.51
683-1	Metal. sed.	18.702	15.531	38.307		0.95
683-5	Metal. sed.	18.713	15.536	38.308		6.42
683-7	Metal. sed.	18.742	15.559	38.379		9.58
683-9	Metal. sed.	18.714	15.568	38.349		15.98
Other sediments						
684-7	Carbonate	18.789	15.553	38.356	49.4	11.56
684-7	Carbonate	18.798	15.570	38.409		11.56
684-8	Carbonate	18.806	15.606	38.557	1.4	13.66
684-8	Carbonate	18.823	15.628	38.617	1.4	13.66
SED 225-27-2 33-35	Evaporite	18.705	15.601	38.175	0.3	
SED 227-40-2 55-57	Black shale	18.989	15.582	38.425	3.0	
SIPAN 2	Detrit. sed.	18.830	15.585	38.467	5.7	
SIPAN 38	Detrit. sed.	18.893	15.593	38.558	4.7	

Lead isotope results are calibrated against SRM 981 measurements, with a mass discrimination factor of 0.1% per atomic mass unit. Typical uncertainties (2σ) are 0.012 for $^{206}\text{Pb}/^{204}\text{Pb}$, 0.014 for $^{207}\text{Pb}/^{204}\text{Pb}$, and 0.03 for $^{208}\text{Pb}/^{204}\text{Pb}$. Pb blanks were routinely measured and were always lower than 0.7 ng.

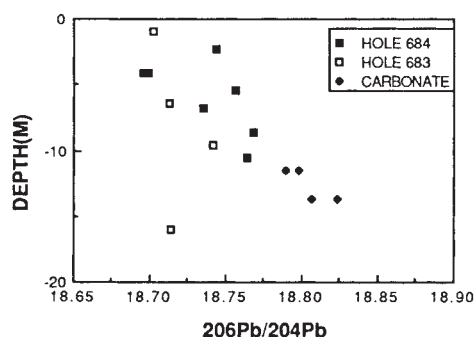


Fig. 3 $^{206}\text{Pb}/^{204}\text{Pb}$ versus depth in cores 684 and 683.

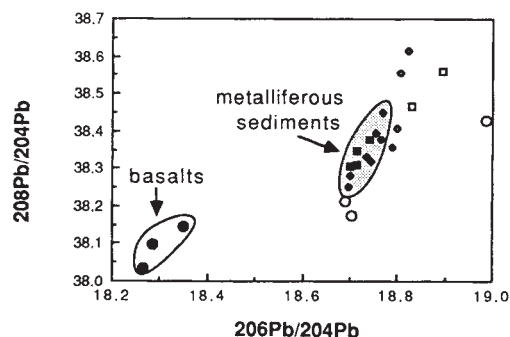


Fig. 4 $^{206}\text{Pb}/^{204}\text{Pb}$ vs $^{208}\text{Pb}/^{204}\text{Pb}$ diagram for Red Sea sediments. Notice that the correlation for the sediments of Atlantis II is well off the local basalts. The symbols are: ■, core 684 sediments; ♦, core 683; □, Red Sea Recent sediments; ◇, carbonate sediments (zone E, core 684); ○, samples in evaporite series (our results and ref. 11).

metals can be extracted from them. Next, the fluid passes through the Miocene evaporite sequence and continues to extract lead. Eventually the fluid may interact with the basalt, but without extracting more lead. Thus, in this hypothesis, fluid circulation would be responsible for the mixing between the sources of lead. (2) The oceanic sedimentation has provided a continuous influx of lead with isotopic compositions similar to the Recent surrounding sediments. Therefore, the lead isotopic composition observed in the metalliferous sediments could result from mixing between the normal sedimentary supply of Pb derived from a fluid that percolated through the evaporites sequence. The rate of input of metal-bearing fluid could vary with time and explain the observed variations. (3) As discussed above, the lead isotopic compositions of the metalliferous sediments vary from the carbonate sediment end-member towards the metalliferous sediment end-member, as determined in core 683, but this trend is discontinuous (Fig. 4). The observed trend in core 684 (W basin) and 683 (SW) is explained if we assume that the fluid was injected first in the SW basin. The brine thickness increased in the SW basin and gradually spilled into the W basin. This mechanism explains the decrease of the isotopic results in the W basin, without requiring an input of fluid directly below this basin. The discontinuous variations of the lead isotopic composition along metalliferous sediments column in the W basin may correspond to interruptions of the fluid input in the SW basin. Taking analytical errors into account, lead isotopic variations occur in the column of the sedimentary SW basin. The most radiogenic value determined in the SW basin (sample 683-7) could also be explained by a hiatus in the input of fluid. During this quiescent period, the lead content could be homogenized by mixing between the brines of the SW and W basins to yield an increased $^{206}\text{Pb}/^{204}\text{Pb}$ ratio for the newly settled sediments. Therefore, the observed isotopic variations in the W and SW basin could be a result of several processes of mixing between the brines of the various Atlantis II basins (North and East

basins included). The similar isotopic values determined near the top of the 683 and 684 cores suggest that an isotopic homogenization occurred in the Recent lower brine of the Atlantis II Deep.

The Pb isotopic compositions of Atlantis II Deep metalliferous sediments are unique when compared to those from ocean ridges. Results from the East Pacific Rise at 21°N (refs. 16, 17), 7°N (ref. 18) and Juan de Fuca Ridge¹⁹ show that the lead isotopic values of sulphides in these areas are similar to the associated basalt values. This result indicates that the oceanic hydrothermal activity efficiently extracts lead from basalts. Nevertheless, a direct comparison should not be made to a spreading centre overlain by sediments (such as the Guaymas basin). In these areas, the lead isotopic compositions of hydrothermal sulphides are either similar to the sediment values, or intermediate between the sedimentary and basaltic values^{20,21}.

We thank the Terres Australes et Antarctiques Françaises for help during the cruise and M. G. Lussereau for some sample analysis; also V. Ragnarsdottir and A. P. Lehuray for useful comments on the manuscript. This work has been supported by CNRS ATP 'Métallogénie'.

Received 13 October 1987; accepted 23 March 1988.

1. Blanc, G., Boulègue, J., Badaut, D. & Stouff, P. C. *acad. Sci., Paris*, **302**, 175-180 (1986).
2. Blanc, G. thesis, Univ. Pierre et Marie Curie (1987).
3. Bäcker, H. & Richter, H. *Geol. Rdsch.* **62**, 697-741 (1973).
4. Shanks, W. L. & Bischoff, J. L. *Econ. Geol.* **75**, 445-459 (1980).
5. *Hot Brines and Recent Heavy Metal Deposits of the Red Sea* (eds Ku, T. L., Turber, D. L. & Mathieu, G. G.) 474-498 (Springer, New York, 1969).
6. Geyh, M. A. & Höndorf, A. *Geol. Jb.* **17**, 79-91 (1976).
7. Manhès, G., Minster, J. F. & Allègre, C. J. *Earth planet. Sci. Lett.* **39**, 37-51 (1978).
8. Dupré, B. & Allègre, C. J. *Nature* **303**, 142-146 (1983).
9. Hamelin, B. & Allègre, C. J. *Nature* **315**, 196-199 (1985).
10. Hamelin, B., Dupré, D. & Allègre, C. J. *Earth planet. Sci. Lett.* **76**, 288-298 (1985).
11. Delevaux, M. H. & Doe, B. R. *Init. Rep. Deep Sea Drilling Project* **23**, 947-950 (US Govt Printing Office, Washington, 1974).
12. *Hot Brines and Recent Heavy Metal Deposits of the Red Sea* (eds Emery, K. O., Hunt, J. M. & Hays, E. E.) 576-571 (Springer, New York, 1969).
13. Hart, R. *Init. Rep. Deep Sea Drilling Project* **34**, 283-335 (US Govt Printing Office, Washington, 1976).
14. Dymond, J., Corliss, J. B. & Stillinger, R. *Init. Rep. Deep Sea Drilling Project* **34**, 575-588 (US Govt Printing Office, Washington, 1976).
15. Dymond, J. in *Nazca plate: Geol. Soc. Am. Mem.* **154**, 133-173 (1981).
16. Brévard, O., Dupré, B. & Allègre, C. J. *Econ. Geol.* **76**, 1205-1210 (1981).
17. Vidal, P. & Clauer, N. *Earth planet. Sci. Lett.* **55**, 237-246 (1981).
18. Boulègue, J., Perseil, E. A., Bernat, M., Francheteau, J. & Dupré, B. *Earth planet. Sci. Lett.* **70**, 549-259 (1984).
19. Hegner, E. & Tatsumoto, M. *J. geophys. Res.* **92**, 11381-11386 (1987).
20. Chen, J. H., Wasserburg, G. J., Von Damm, K. L. & Edmond, J. M. *Geochim. cosmochim. Acta* **50**, 2567-2579 (1986).
21. Lehuray, A. P., Church, S. E., Koski, R. A. & Bouse, R. M. *Geology* (in the press).

Carbon isotope compositions of fluid inclusions in charnockites from southern India

D. H. Jackson, D. P. Matthey & N. B. W. Harris

Department of Earth Sciences, The Open University,
Milton Keynes MK7 6AA, UK

Models for the formation of crustal granulites, which include the metamorphism of anhydrous lithologies^{1,2}, the extraction of hygroscopic granitic melts^{3,4} and fluxing by CO_2 ⁵⁻⁷, constrain the thermal and geochemical evolution of the lower crust and its interaction with mantle-derived volatiles. Here we present abundances and carbon isotope compositions of CO_2 in fluid inclusions in quartz and garnet from gneiss-granulite pairs from the incipient charnockite quarries of southern India, determined by a stepped combustion/thermal decrepitation carbon-extraction technique, which quantifies problematic sample contamination and allows precise lithological control. The results show that the CO_2 in incipient charnockite is more abundant and isotopically heavier than the

CO₂ in associated gneiss, providing evidence for the influx of CO₂-rich fluids during their formation. The inferred carbon isotope compositions of these fluids are consistent with a mantle source, and at Ponmudi, the isotopic composition of coexisting graphite and CO₂ implies fluid/rock volume ratios in excess of 0.15.

Incipient charnockites from southern India, which occur as veins and patches within biotite and amphibole gneiss, argue for a volatile-controlled transition from amphibolite-facies gneiss to arrested charnockite development⁸⁻¹¹. The presence of grain-margin calcite⁸ and abundant CO₂-rich fluid inclusions within the charnockite^{12,13} have been invoked as evidence for the presence of a CO₂-rich fluid during metamorphism. Whether this CO₂ is from an external mantle or lower-crustal source⁸, or derived internally by oxidation of graphite¹² or pore-fluid release during uplift¹⁴ remains controversial.

The character of fluid inclusions in incipient charnockites varies across the Palghat-Cauvery (P-C) shear zone, one of a system of late Proterozoic shear belts that transect the granulite terranes (Fig. 1)¹⁵. North of the P-C shear zone, high-density CO₂-rich inclusions trapped in quartz are post-dated by smaller water-rich inclusions¹². Entrapment pressures of the high-density inclusions calculated from the CO₂ equation of state are within 1 kbar of the peak metamorphic pressure (5-6 kbar)¹². South of the shear zone, early inclusions are seen to be CO₂-rich, again trapped close to the peak of metamorphism, but a younger generation of mixed CO₂-H₂O inclusions was trapped at much lower pressures¹³.

In this study, seven pairs of incipient charnockites and their gneissic precursors (normally sampled between 10 cm and 1 m apart) from both north and south of the P-C shear zone have been selected for carbon isotope analysis (Table 1). CO₂ was liberated from fluid inclusions in separated quartz or garnet with a stepped combustion/thermal decrepitation technique¹⁶. This procedure permits samples as small as individual mineral grains (60-120 mg) to be analysed, compared with 1-50 g required by previous studies^{5,17,18}. This ability is critical to studies of incipient charnockites, which require precise lithological control of closely associated amphibolite- and granulite-facies material. Each mineral separate was ultrasonically cleaned, degassed to 10⁻⁵ torr, and combusted at 400 °C in pure

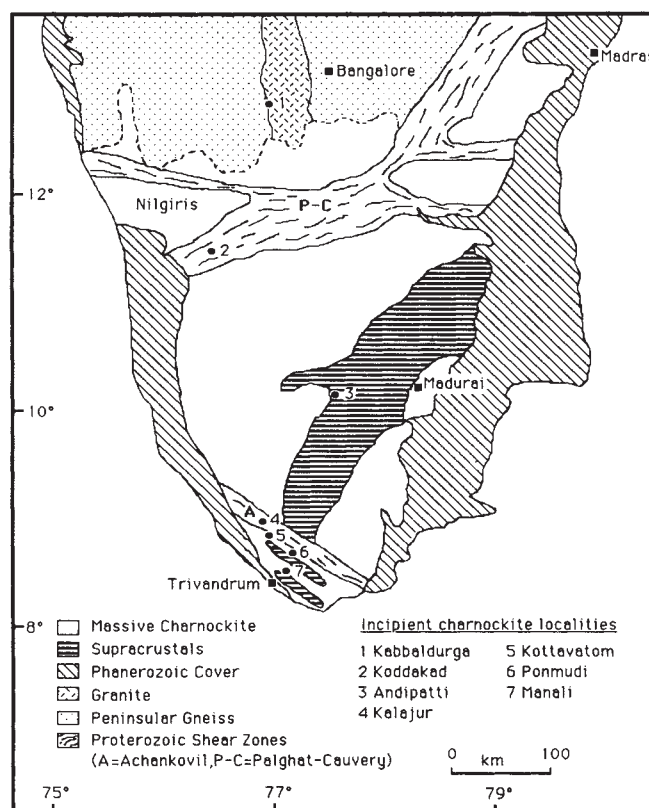


Fig. 1 Simplified geological map of southern India, showing sample localities.

oxygen to remove residual surface contamination¹⁹. The sample was then pyrolysed for 30 min in 100 °C steps up to 1,200 °C. The isotopic composition of CO₂ released from each step was analysed by a SIRA 24 mass spectrometer to a precision typically better than ±0.1%. System blanks (<20 ng C per 100 °C step above 400 °C) were low enough to be ignored.

Similar carbon release profiles were obtained for nearly all the samples in this study (see Fig. 2 inset). Carbon released by the combustion step at 400 °C is isotopically light ($\delta^{13}\text{C}$ values (see Table 1) ranging from -30‰ to -25‰) and represents adsorbed organic contamination¹⁹. With increasing temperature, carbon yields sharply increase to form a single maximum release of isotopically heavier CO₂ in the 600-800 °C range (Fig. 2 inset). Above 800 °C carbon yields diminish rapidly, with essentially all the CO₂ released by 1,000 °C. The integrated abundance and isotopic composition of CO₂ released by pyrolysis across the 600-800 °C 'peak release' interval are summarized in Table 1. Optical observations of heating-stage decrepitations on several incipient charnockites and precursor gneisses from Kerala confirm that the high-density CO₂-rich inclusions rupture in the same temperature range as the 'peak release' of CO₂ determined in this study²⁰. We therefore interpret gas release in this temperature range as CO₂ liberated by the decrepitation of fluid inclusions, as concluded in fluid inclusion studies of granulites elsewhere¹⁸.

The abundance and isotopic composition of CO₂ from fluid inclusions in incipient charnockite-gneiss pairs are plotted in Fig. 2. In all but one case (Andipatti, where the two lithologies were taken from only 1 cm apart), sample pairs clearly indicate a substantial enrichment (up to a factor of four) of CO₂ in the charnockite relative to the associated gneiss. Furthermore, the $\delta^{13}\text{C}$ of quartz-bound fluid-inclusion CO₂ is up to 4‰ heavier than CO₂ from the associated gneiss. These data suggest that a CO₂-rich fluid phase containing isotopically heavier carbon than

Table 1 Abundance and isotopic composition of carbon in gneiss/incipient charnockite pairs

Locality	Sample no.	Lithology	Mineral	C (p.p.m.)	$\delta^{13}\text{C}^*$
1 Kabbaldurga	KA2	Gneiss	Quartz	7.7	-13.3
	KA4	Charnockite	Quartz	16.5	-11.9
	KA5	Charnockite	Quartz	18.8	-11.4
2 Koddakad	PC6A	Gneiss	Quartz	11.7	-7.1
	PC6C	Charnockite	Quartz	43.8	-7.0
3 Andipatti	AN11A1	Gneiss	Quartz	39.8	-6.3
	AN11A2	Charnockite	Quartz	38.3	-6.3
	TR10A	Gneiss	Quartz	32.0	-7.7
4 Kalajur	TR10F	Gneiss	Quartz	44.6	-7.1
	TR10D	Charnockite	Quartz	66.9	-6.5
	SH/QF12§	Gneiss	Quartz	12.0	-11.0
5 Kottavatom†	SH/QF/11§	Charnockite	Quartz	42.2	-10.0
	SH/QF/3B§	Gneiss	Quartz	22.4	-11.2
6 Ponmudi†	TR15A	Gneiss	Garnet	19.2	-12.6
	TR15A	Gneiss	Graphite	—	-19.7
	SH/QF/3§	Charnockite	Quartz	76.7	-10.4
	SH/QF/3§	Charnockite	Quartz	50.4	-10.3
	TR15B	Charnockite	Quartz	66.4	-9.1
	TR15B	Charnockite	Garnet	70.2	-8.5
	TR15B	Charnockite	Graphite	—	-18.1
7 Manali‡	SH/QF/6§	Gneiss	Quartz	44.6	-11.7
	SH/QF/5§	Charnockite	Quartz	150.2	-7.6

Numbered localities from Fig. 1. Data for quartz and garnet from CO₂ released over the 600-800 °C peak-release temperature interval.

* $\delta^{13}\text{C} = [((^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}}) - 1] \times 1000$. The standard is PDB carbonate.

† Graphite-bearing.

‡ Quartz segregations in graphite-bearing charnockite.

§ Data from ref. 20.

the precursor gneiss was associated with the formation of the charnockite.

The graphite present in samples from south of the Achankovil shear zone is a potential source of CO_2 for the charnockitization process. At Ponmudi, the $\delta^{13}\text{C}$ of graphite is -19.7% in the gneiss and -18.1% in the charnockite (Table 1). A quartz-graphite fractionation in both gneiss and charnockite of $\Delta\delta^{13}\text{C}_{\text{CO}_2(\text{qtz})-\text{graphite}} \approx 8\%$ suggests an equilibration temperature of $\sim 700^\circ\text{C}$ ²¹, which compares well with peak temperatures of $700\text{--}800^\circ\text{C}$ determined from phase equilibria^{12,14}, confirming CO_2 entrapment near peak metamorphic conditions and arguing against a post-granulite phase of fluid influx as has been documented in granulites from the Adirondacks²². The quartz and garnet in precursor gneiss contain relatively low abundances of CO_2 , with $\delta^{13}\text{C}$ values of -11.2% to -12.6% (Table 1). Whether this CO_2 is wholly derived internally from graphite oxidation, or represents a more complex mixture of pre- and syn-metamorphic components is open to question, but it is clear that if graphite oxidation were a significant process then the $\delta^{13}\text{C}$ of CO_2 -enriched charnockite would be expected to approach that of the graphite. Instead, an increase of $\approx 1.5\%$ in $\delta^{13}\text{C}$ in charnockite relative to gneiss, recorded by both graphite and the average composition of quartz-bound CO_2 , implies that the enrichment of CO_2 in the charnockite was caused by influx of CO_2 whose $\delta^{13}\text{C}$ was heavier than in the precursor gneiss. Although graphite is precluded as a significant source of CO_2 at this locality, it is clear from Fig. 2 that, excluding Kabbaldurga, the $\delta^{13}\text{C}$ of fluid-inclusion carbon is lighter in graphite-bearing than in graphite-free localities. We interpret this as evidence for equilibration between graphite (usually present at modal concentrations of $\approx 1\%$ ¹²) and influxed carbon.

The isotopic shift in graphite and CO_2 constrains the fluid/rock volume ratio¹⁷. At Ponmudi, a greater contrast in $\delta^{13}\text{C}$ values between the gneiss and charnockite is recorded by fluid inclusions in garnet (4.1%) than by inclusions in quartz ($\approx 1.5\%$, Table 1). As garnet is partially consumed in the charnockite-forming reactions¹², these results suggest that whereas quartz-bound CO_2 and graphite were able to equilibrate during charnockite formation, the $\delta^{13}\text{C}$ of CO_2 in garnet-bound inclusions (-8.5%), having been retained early in the granulite event, approaches the composition of the influxed fluid. Assuming that the measured $\delta^{13}\text{C}$ of -19.7% of graphite in Ponmudi gneiss represents the initial graphite composition and $\Delta\delta^{13}\text{C}_{\text{CO}_2(\text{qtz})-\text{graphite}} = 8\%$, the 1.5% shift in the $\delta^{13}\text{C}$ of charnockitic graphite requires a fluid/rock volume ratio of ~ 0.15 . This value is a minimum constraint and would increase if the amount of graphite in the precursor gneiss exceeded 1% , or if the initial graphite $\delta^{13}\text{C}$ were lighter than the observed value of -19.7% . A fluid/rock ratio of this value in granulites is high, and compares with similar values calculated from metasediments exposed to circulating CO_2 -rich fluids during amphibolite-grade metamorphism at Naxos^{17,23}.

$\delta^{13}\text{C}$ values of mid-ocean-ridge basalt glasses, carbonatites and diamonds, generally indicative of mantle carbon, have a range of values with a well-defined peak between -5% and -8% ²⁴. The inferred $\delta^{13}\text{C}$ of influxed fluid at Ponmudi lies close to -8.5% , and $\delta^{13}\text{C}$ values of CO_2 from graphite-free incipient charnockites in the southern group lie between -6% and -7% . These data are consistent with a mantle source for influxed CO_2 , although where graphite is present its isotopic composition is buffered to lighter values (-7.6% to -10.0%). Gneisses and charnockites from the graphite-free Kabbaldurga quarry north of the P-C shear zone are characterized by lighter $\delta^{13}\text{C}$ than graphite-free samples found to the south, although the shift to heavier carbon between gneiss and charnockite is retained. These data also imply an external source of CO_2 , with $\delta^{13}\text{C}$ heavier than that of the precursor gneiss, but whether this influxed CO_2 has a mantle isotopic signature is more equivocal. Different mechanisms for incipient charnockite formation at Kabbaldurga and in localities south of the P-C shear zone have been proposed

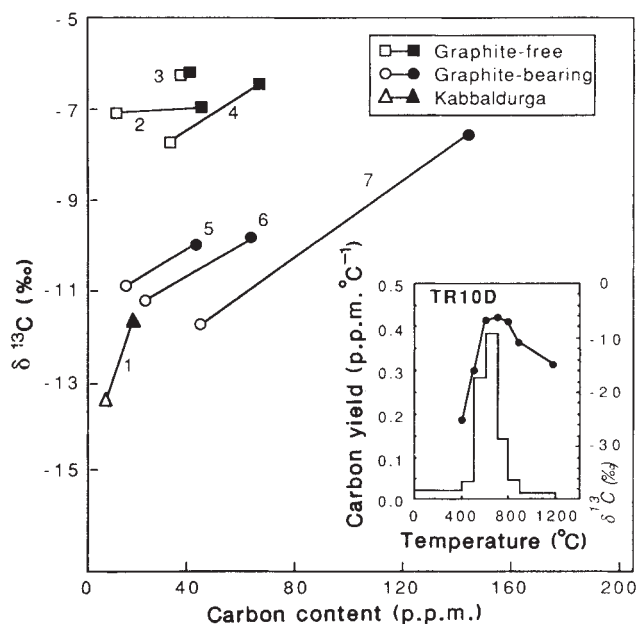


Fig. 2 Mean abundances and isotopic compositions of carbon released from quartz across the $600\text{--}800^\circ\text{C}$ temperature interval in gneiss (open symbols) and associated incipient charnockite (closed symbols). Locality numbers from Fig. 1. Inset, typical stepped-heating profile for charnockitic quartz from Kalajur, showing the abundance (histogram) and isotopic composition (closed circles) of carbon released (as CO_2) at each temperature step.

on the basis of contrasting mineral reactions and fluid-inclusion characteristics^{12,14}. The data presented here substantiate this difference, yet suggest that influx of CO_2 -rich fluids from an external source was an important mechanism in both areas. Granulites from the northern terrane provide mineral and whole-rock radiometric ages of $\approx 2,500\text{ Myr}$ ²⁵⁻²⁷, whereas to the south, radiometric ages indicate a much younger period of granulite metamorphism of $\approx 550\text{ Myr}$ ²⁸, roughly synchronous with shear-zone formation and carbonatite emplacement within the shear zones^{29,30}. Clearly, granulite formation may have occurred during more than one period in the evolution of the southern Indian crust, but the carbon isotope data from throughout southern India argue for an external source of CO_2 during the equilibration of the incipient charnockite assemblages.

We gratefully acknowledge Dr V. Subramanian and the Department of Mining and Geology in Madras, Dr S. Jayaram and the Department of Mines and Geology in Bangalore and Dr M. Santosh and the Centre for Earth Science Studies in Trivandrum for extensive support in the collection of samples for this study and Colin Pillinger for use of stable isotope facilities at the OU.

Received 18 December 1987; accepted 2 March 1988.

- Lamb, W. & Valley, J. W. *Nature* **312**, 56-58 (1984).
- Martignole, J. *Precamb. Res.* **9**, 303-310 (1979).
- Nesbitt, H. W. *Contr. Miner. Petrol.* **72**, 303-310 (1980).
- Pride, C. & Muecke, G. K. *Contr. Miner. Petrol.* **73**, 403-412 (1980).
- Hoefs, J. & Touret, J. *Contr. Miner. Petrol.* **52**, 165-174 (1975).
- Collerson, K. D. & Fryer, B. S. *Contr. Miner. Petrol.* **67**, 151-167 (1978).
- Newton, R. C., Smith, J. V. & Windley, B. F. *Nature* **288**, 45-40 (1980).
- Hansen, E. C., Jarnardhan, A. S., Newton, R. C. & Smith, J. V. *Nature* **278**, 511-517 (1979).
- Friend, C. R. L. *Nature* **294**, 550-552 (1981).
- Newton, R. C. & Hansen, E. C. in *The Nature of the Lower Crust* (eds Dawson, J. B., Carswell, D. A., Hall, J. & Wedepohl, K. H.) *Geol. Soc. spec. Publ.* **24**, 297-307 (1986).
- Stahle, H. J., Raith, M., Hoernes, S. & Delfs, S. *J. Petrology* **28**, 803-834 (1987).
- Hansen, E. C., Jarnardhan, A. S., Newton, R. C., Prame, W. K. B. N. & Ravindra Kumar, G. R. *Contr. Miner. Petrol.* **96**, 225-244 (1987).
- Santosh, M. *Lithos* **19**, 1-10 (1986).
- Raith, M., Hoernes, S., Klatt, E. & Stahle, H. J. in *Fluid movements, Element Transport and the Composition of the Deep Crust* (ed. Bridgewater, D.) (Reidel, Dordrecht, in the press).
- Drury, S. A., Harris, N. B. W., Holt, R. W., Reeves-Smith, G. J. & Wightman, R. T. *J. Geol.* **92**, 3-20 (1984).
- Swart, P. K., Grady, M. M. & Pillinger, C. T. *Meteoritics* **18**, 137-154 (1983).
- Kreulen, R. *Contr. Miner. Petrol.* **92**, 28-32 (1988).

18. Pineau, F., Javoy, M., Behar, F. and Touret, J. *Bull. minéral.* **104**, 630-641 (1981).
19. DesMarais, D. *Earth planet. Sci. Lett.* **79**, 21-26 (1986).
20. Santosh, M., Jackson, D. H., Matthey, D. P. & Harris, N. B. W. *J. geol. Soc. India* (in the press).
21. Bottinga, Y. *Geochim. cosmochim. Acta* **3**, 49-64 (1969).
22. Lamb, W., Valley, J. W. & Brown, P. E. *Contr. Miner. Petrol.* **96**, 485-495 (1987).
23. Schuling, R. D. & Kreulen, R. *Earth planet. Sci. Lett.* **43**, 298-302 (1979).
24. Matthey, D. P. *Terra Cogn.* **7**, 31-37 (1987).
25. Spooner, C. M. & Fairbairn, H. W. *J. geophys. Res.* **75**, 6707-6713 (1970).
26. Grew, E. S. & Manton, W. J. *J. geol. Soc. India* **25**, 193-195 (1984).
27. Buhl, D., Grauert, B. & Raith, M. *Fortschr. Miner.* **61**, 43-45 (1983).
28. Buhl, D. thesis, Univ. Munster (1987).
29. Deans, T. & Powell, J. L. *Nature* **218**, 750-752 (1968).
30. Subramanian, V. thesis, Indian Inst. Sci., Bangalore (1983).

Isolation of insect juvenile hormone III from a plant

Yock C. Toong*, David A. Schooley† & Fred. C. Baker†

* Universiti Sains Malaysia, School of Chemical Sciences, 11800 Penang, Malaysia

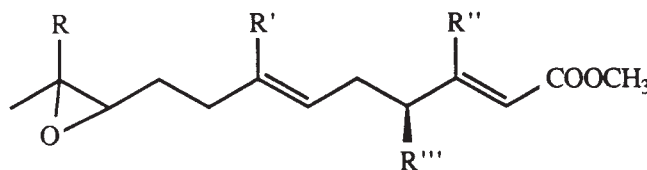
† Sandoz Crop Protection, Zoecon Research Institute, 975 California Avenue, Palo Alto, California 94304, USA

Juvenile hormones (JHs) are insect hormones which promote the retention of larval structures when present at larval-larval moults, but which also stimulate ovarian growth and maturation in adult insects. Here we report the isolation of JH III from the Malaysian plant *Cyperus iria* (common name grasshopper's Cyperus; family, Cyperaceae) along with the closely related compound, methyl (2*E*,6*E*)-farnesoate. The latter is a putative JH of crustaceans¹ and some insect species². The plant JH III has the same 10*R* configuration as that previously determined for JH III secreted by insect tissue³. Grasshopper (*Melanoplus sanguinipes*) nymphs fed on *C. iria* grew at a rate similar to controls fed on wheat seedlings, but showed pronounced morphogenetic effects due to excess JH (metathetely⁴) on moulting to adults. Ovaries dissected from *Cyperus*-fed females were markedly underdeveloped compared to those of normal females. These findings demonstrate a novel plant defence mechanism against insects.

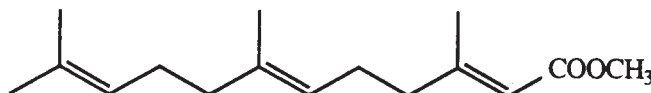
The juvenile hormones are a family of five closely related sesquiterpenoid compounds⁵ (Fig. 1) important in the regulation of insect development and reproduction. These hormones are secreted by a pair of minute retrocerebral organs, the corpora allata (CA). Analysis of miscellaneous insect species has shown the structurally simplest member, JH III, to be the most ubiquitous of the JHs⁶. Recently, the unepoxidized precursor of JH III, methyl farnesoate (Fig. 1), was isolated from the haemolymph of the crustacean *Libinia emarginata*^{1,7} and from *Homarus americanus*⁷ and from embryos of the cockroach *Nauphoeta cinerea*². In addition, methyl farnesoate is secreted *in vitro* by embryonic *N. cinerea* corpora cardiaca-corpora allata² and by crustacean mandibular organs⁷, glands thought to be analogous to the insect CA. This compound may function as a JH or pro-JH in crustaceans^{1,7} and insects².

Since the discovery of the JHs there has been much interest in the synthesis of analogues for use as insect control agents, and several of these have enjoyed commercial success⁸. Many natural products, including plant and animal extracts, have been screened for JH activity on miscellaneous insect species. An early example was the detection of the 'paper-factor' from Balsam fir⁹, subsequently isolated and named juvabione¹⁰. Additional natural products with JH activity have subsequently been isolated and identified from plants^{11,12}. None of these compounds closely resembles the JHs in chemical structure; furthermore, most plant-derived juvenoids are active only on a narrow range of insect species.

Whole plants of *C. iria*, collected in the field in Penang, Malaysia, were chopped finely and steam distilled. The distillate was extracted with chloroform and the concentrated chloroform



R=R'=R''=Me,	R'''=H,	(10 <i>R</i>)-JH III
R=Et, R'=R''=Me,	R'''=H,	JH II
R=R'=Et, R''=Me,	R'''=H,	(10 <i>R</i> ,11 <i>S</i>)-JH I
R=R'=R''=Et,	R'''=H,	JH O
R=R'=Et, R''=Me,	R'''=Me,	4-MeJH I



Methyl Farnesoate

Fig. 1 Structures of the known JHs and methyl (2*E*,6*E*) farnesoate.

layer was purified by preparative TLC (silica gel developed with 10% ethyl acetate in 60-80°C petroleum ether). A chromophoric zone with $R_f \approx 0.3-0.4$ was scraped and eluted with chloroform to give an optically active oil, $[\alpha]_{578}^{25} + 4.92^\circ$ methanol (c 0.502). Analysis by NMR and mass spectrometry gave spectral data consistent with a mixture containing JH III and the eudesmanoid, intermedeol. While a positive optical rotation value suggested the presence of (+)-(10*R*)-JH III, contamination of the latter with (+)-intermedeol¹³ ($[\alpha]_{578}^{25} + 10.7$) made this result equivocal. The mixture was further purified by C_{18} Sep-Pak and by silica Sep-Pak filtration followed by sequential preparative silica adsorption liquid chromatography (LC) (two different solvent systems), and reversed-phase (C_8) LC. Analysis by silica adsorption LC (detection at 254 nm) indicated that the product was 99.5% pure. The electron impact mass spectrum of the product was identical to that from a synthetic sample of (10*RS*)-JH III.

To measure the chirality of the natural product, we converted an aliquot (containing a trace of (*RS*)-[10-³H]JH III) to the corresponding JH III 10,11-diol¹⁴ (Fig. 2) using 0.1 M H_2SO_4 in $H_2^{18}O$ (>99% ^{18}O). Electron impact mass spectral fragmentation of unenriched JH III 10,11-diol produces the important ions m/z 59 and 225 from cleavage of the C-10, -11 bond (Fig. 2). Enrichment with ^{18}O at the C-11 or C-10 hydroxyl group shifts these ions to m/z 61 or 227, respectively. We analysed the oxygen-18 enriched JH III diol by selected ion monitoring gas chromatography/mass spectrometry: 99.5% of the ^{18}O was located at the C-11 position (m/z 61) and only 0.5% at the C-10 position (m/z 227) of the JH III 10,11-diol. As attack by H_2O at C-10 versus C-11 gives the diol of the opposite absolute configuration¹⁴, JH III was converted to the diol with minimal loss of chirality. The 10,11 diol was reacted with (*R*)- α -methoxy- α -trifluoromethylphenylacetyl chloride¹⁵ (MTPC1, >99.7% (*R*)) in pyridine to form the corresponding JH III 11-OH,10-(*R*)-MTP ester³. The diastereomeric esters completely separate on normal phase LC. The major UV absorbing product (97.3%) was the faster eluting (10*R*, αR)-diastereomer³, while only ~2.7% 10*S*, αR diastereomer was observed. Thus, JH III isolated from the plant *C. iria* has the same 10*R* absolute configuration as insect JH III³. Liquid scintillation counting of LC eluent indicated that label from (*RS*)-[10-³H]JH III tracer was evenly distributed between the 10*R* and 10*S* diastereomers as expected. JH III isolated from *C. aromaticus* collected in the field also possesses the 10*R* configuration.

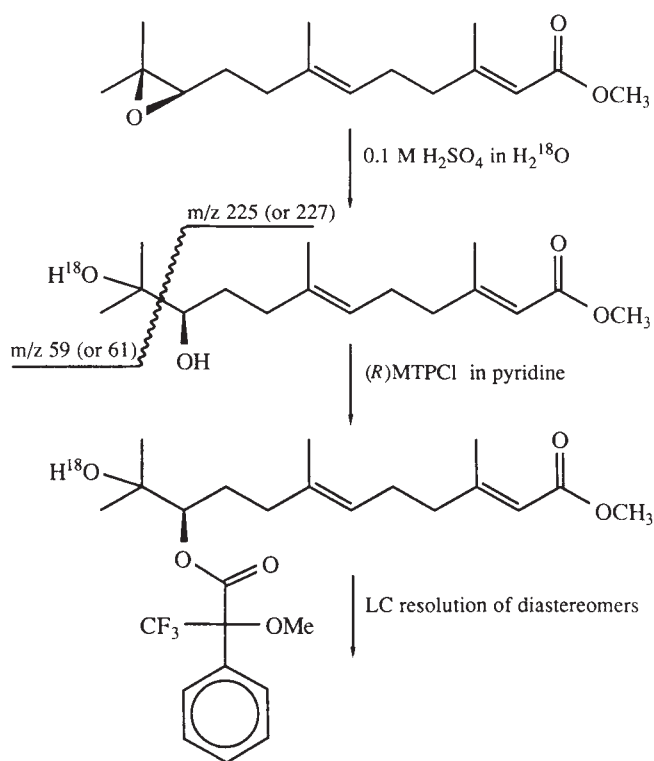


Fig. 2 Scheme for analysis of chirality of plant JH III. Hydration of the epoxide with acidic H₂¹⁸O gave a diol containing 99.5% of oxygen-18 at C-11. Reaction of the diol with the acid chloride derived from (+) MTPA (believed to have the *R* configuration¹³) gave diastereomeric esters. These have a known elution order³ on LC.

Cyperus iria was subsequently grown from seed under greenhouse conditions and mature plants (3 months old) were processed to determine the JH titre¹⁶. Levels of JH III (detected as the *d*₃-methoxyhydrin derivative) were so high ($\approx 151 \mu\text{g g}^{-1}$ fresh weight), that quantification was achieved simply using LC analysis and integrating the UV absorbing zone corresponding in retention time to that of an authentic standard. None of the other known JHs could be detected by a mass spectrometric method¹⁶ with a limit of detection of $\approx 0.05 \text{ ng g}^{-1}$. For comparison, JH levels detected in most insect species are in the range 0.1–100 ng per g fresh weight, although a few species contain up to $\approx 1 \mu\text{g g}^{-1}$ at specific stages during development. We also detected significant levels ($\approx 14 \mu\text{g per g}$ fresh weight) of methyl (2*E*, 6*E*)-farnesoate in *C. iria* samples. The electron impact mass spectrum of the product was identical to that of authentic standard.

Third stadium grasshopper nymphs (*M. sanguinipes*) fed on flowering *C. iria* plants ate normally and continued to grow and moult just like control insects during the two subsequent nymphal instars. Following the final moult, however, 90% of the adults were abnormal compared to control animals reared on wheat seedlings. Typical metathetic effects⁴ were observed, including twisted wings and colour changes, reminiscent of treatment with excess JH. In addition, the ovaries from females raised on *C. iria* contained markedly underdeveloped eggs compared to controls and these females laid no egg pods.

The detection of very high levels of JH III in *C. iria* (and *C. aromaticus*) and the demonstration of clear morphogenetic effects on insects raises some intriguing ecological implications for these plants in their native habitat. Clearly, *C. iria* and other species represent a valuable source of (10*R*)-JH III for research purposes.

We thank Ms Leslie W. Tsai, Dr Gerardus B. Staal, Ms C.

Reuter, Mr G. Jamieson, Dr H. Röller and Dr G. B. Quistad for assistance and/or advice and the NSF for partial financial support.

Received 23 February; accepted 7 April 1988.

1. Laufer, H. *et al. Science* **235**, 202–205 (1987).
2. Lanzrein, B., Imboden, H., Bürgin, C., Brüning, E. & Gfeller, H. in *Biosynthesis, Metabolism and Mode of Action of Invertebrate Hormones* (eds Hoffman, J. & Porchet, M.) 454–465 (Springer, Berlin, 1984).
3. Judy, K. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **70**, 1509–1513 (1973).
4. Novak, V. J. A. *Insect Hormones* (Methuen, London, 1966).
5. Schooley, D. A. & Baker, F. C. in *Comprehensive Insect Physiology Biochemistry and Pharmacology* Vol. 7 (eds Kerkut, G. A. & Gilbert, L. I.) 363–389 (Pergamon, Oxford, 1985).
6. Schooley, D. A., Baker, F. C., Tsai, L. W., Miller, C. A. & Jamieson, G. C. in *Biosynthesis, Metabolism and Mode of Action of Invertebrate Hormones* (eds Hoffmann, J. & Porchet, M.) 373–383 (Springer, Berlin, 1984).
7. Borst, D. W. *et al. Insect Biochem.* **17**, 1123–1127 (1987).
8. Henrick, C. A. in *Insecticide Mode of Action* (ed. Coats, J. R.) 315–402 (Academic, New York, 1982).
9. Slama, K. & Williams, C. M. *Proc. natn. Acad. Sci. U.S.A.* **54**, 411–414 (1965).
10. Bowers, W. S., Fales, H. M., Thompson, M. J. & Uebel, E. C. *Science* **154**, 1020–1021 (1966).
11. Černý, V., Dolejš, L., Labler, L., Sorm, F. & Slama, K. *Tetrahedron Lett.* 1053–1057 (1967).
12. Nishida, R., Bowers, W. S. & Evans, P. H. *Arch. Insect Biochem. Physiol.* **1**, 17–24 (1983).
13. Zalkow, L. H., Zalkow, V. B. & Brannon, D. R. *Chem. Ind. Černý* **38** (1963).
14. Nakanishi, K., Schooley, D. A., Koreeda, M. & Dillon, J. *JCS chem. Commun.* 1235–1236 (1971).
15. Dale, J. A., Dull, D. L. & Mosher, H. S. *J. org. Chem.* **34**, 2543–2549 (1969).
16. Bergot, B. J., Ratcliff, M. & Schooley, D. A. *J. Chromat.* **204**, 231–244 (1981).

Long-term heterosynaptic inhibition in *Aplysia*

P. G. Montarolo*, E. R. Kandel & S. Schacher

Howard Hughes Medical Institute and Center for Neurobiology and Behavior, Columbia University, College of Physicians and Surgeons and New York State Psychiatric Institute, 722 West 168 Street, New York, New York 10032, USA

Synaptic transmission between mechanosensory and motor neurons of the gill withdrawal reflex in *Aplysia* can undergo both short-term^{1,2} and long-term^{3,4} modulation. One form of short-term synaptic depression lasting minutes can be evoked by the peptide Phe-Met-Arg-Phe-amide (FMRFamide)^{5–8}, and is mediated by the lipooxygenase pathway of arachidonic acid⁷. We report here using cell culture⁹, that the same monosynaptic sensory-to-motor component of the gill withdrawal reflex can also undergo long-term synaptic depression lasting 24 h after five applications of FMRFamide over a 2-h period. The long-term depression evoked by FMRFamide is transmitter-specific. Dopamine or low-frequency stimulation of sensory neurons, which also produce short-lasting synaptic depression *in vivo*^{1,5}, failed to evoke a long-term change. As is the case for long-term presynaptic facilitation of this connection with serotonin¹⁰, the long-term depression, but not the short-term, can be blocked when applications of FMRFamide are given in the presence of anisomycin, a reversible inhibitor of protein synthesis. Thus, heterosynaptic depression parallels heterosynaptic facilitation in having a long-term as well as a short-term form, and in both cases the long-term modulation requires the synthesis of gene products not essential for the short-term changes.

An interesting feature of the synapses of sensory neurons in *Aplysia* is that their efficacy can be modulated up or down by specific neurotransmitters. Short-term heterosynaptic facilitation is evoked by activity in identified serotonergic neurons or by a single application of serotonin (5-HT)^{9,11,12}. By contrast, short-term depression can be produced by activity in identified neurons containing FMRFamide immunoreactivity or by a single application of the peptide FMRFamide^{5–8}. We have simplified the neural circuit of the gill withdrawal reflex by isolating the sensory-to-motor synapse in dissociated cell culture to

* Present address: Instituto di Fisiologia Umana, C. so Raffaello 30, 10125 Torino, Italy.

Fig. 1 Repeated applications of FMRFamide evokes depression of sensory-L7 connections lasting 24 h. **a**, The first evoked EPSP in the L7 motoneuron (0-h column) is compared to the first evoked EPSP for the same connection 24 h later (24-h column). Whereas control (1) or dopamine (3) treatments resulted in slight changes in the EPSP, FMRFamide applications (2) evoked a significant depression in the amplitude of the EPSP. **b**, Summary of long-term depression induced by FMRFamide. The height of each bar is the mean $\pm$ s.e.m. of the per cent change in the amplitude of the EPSP retested 24 h after treatment. An overall analysis of variance indicated a difference with treatment ($F = 4.636$; $d.f. = 2, 21$; $P < 0.03$). A comparison of the means (Dunnett's test) indicated that FMRFamide treatment significantly decreased EPSP amplitude compared with the control ($t = 2.72$; $P < 0.05$) and to the dopamine treatment ($t = 2.545$; $P < 0.05$), whereas the dopamine-treated group did not differ from the control.

Methods. Cultures, consisting of one motor cell L7 and one to three sensory neurons, were prepared and maintained for four or five days with methods described previously⁹. During each recording session, the motor cell was impaled with a glass microelectrode (10–15 M Ω) filled with 2.5 M KCl and held at a potential of -30 mV below the resting level to permit accurate measurement of the amplitude of the EPSP. The recording procedure was modified in two ways from those described previously^{9,10,13}. First, sensory cells were stimulated by a brief (0.2–0.5 ms) depolarizing pulse using an extracellular electrode filled with perfusion medium. Second, the cultures were perfused with a 1:1 mixture of L15 (Flow Laboratories) and artificial seawater containing 5% *Aplysia* haemolymph and maintained 18 °C and 19 °C. For one of the sensory neurons, short-term depression produced by the first application of either 0.1 μ M FMRFamide or 1 μ M dopamine was assessed as described in Fig. 3. Untreated control cells received the same number of stimuli. With the washout of the first transmitter application, the electrodes were removed and subsequent 5-min applications of transmitters were given without additional recording. Each application of transmitter was followed by a 25-min washout with perfusion medium. The cells were returned to culture medium and retested 24 h later with the same procedures.

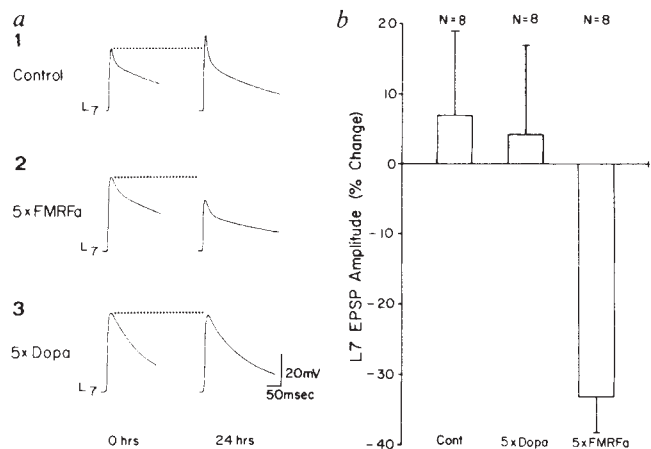
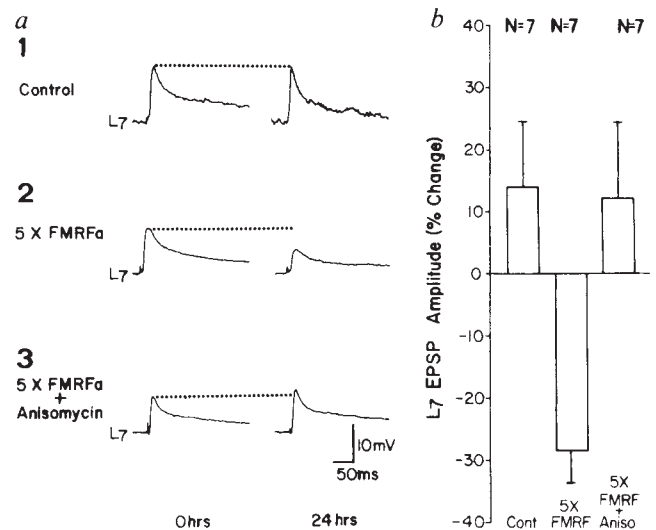


Fig. 2 FMRFamide-induced long-term depression of sensory-L7 connections is blocked in the presence of protein synthesis inhibitor. The protocol used was the same as that described in Fig. 1. Anisomycin (10 μ M) was added to the cultures after the recording, 45 min before the initial FMRFamide application, and was present continuously until 30 min after the last FMRFamide application. **a**, The first evoked EPSP in L7 (0-h column) is compared to the EPSP of the same connections retested 24 h later (24-h column). As was found in the previous experiment (Fig. 1), the control culture showed little change (1) whereas FMRFamide applications produced a marked depression in the amplitude of the EPSP (2). The presence of anisomycin throughout the FMRFamide treatment blocked the decrease of the EPSP amplitude (3). **b**, Summary of the anisomycin blockade of FMRFamide long-term depression. The height of each bar is the same measurement as described in Fig. 1b. An overall analysis of variance indicated a difference with treatment ($F = 6.05$; $d.f. = 2, 18$; $P < 0.01$). A comparison of the means (Dunnett's test) indicated that there is no significant difference between the control group and the FMRFamide treated cultures given in the presence of anisomycin, whereas the repeated FMRFamide applications alone significantly decreased EPSP amplitudes compared to the controls ($t = 3.05$; $P < 0.05$) and the anisomycin treated cultures ($t = 2.943$; $P < 0.05$).



examine the mechanisms whereby short-term synaptic plasticity can be converted to a long-term form. Previously, we used this *in vitro* system to study heterosynaptic facilitation related to behavioural sensitization of the reflex, and found that applying 5-HT once leads to presynaptic facilitation of transmitter release that lasts minutes, whereas five repeated applications result in facilitation lasting 24 h (refs 10, 13). Although 5-HT can produce both short- and long-term facilitation, the long-term change differs from the short-term in requiring macromolecular synthesis¹⁰. We have now extended this analysis to heterosynaptic depression. As in the intact ganglion⁵, the amplitude of the excitatory postsynaptic potential (EPSP) in motor cell L7 evoked by stimulation of the sensory neuron can be depressed for several minutes in response to a single application of dopamine or FMRFamide¹⁴. We therefore examined whether repeated application of dopamine or FMRFamide gives rise to long-term synaptic depression.

We measured the amplitude of the EPSP in L7 evoked by

stimulation of each sensory neuron before (Fig. 1a, 0-h column) and 24 h after (Fig. 1a, 24-h column) the cultures were treated with five applications (5 min each) of either dopamine (1 μ M) or FMRFamide (0.1 μ M) at 30-min intervals. Treatment with FMRFamide caused a significant decrease in the amplitude of the EPSP of $-34 \pm 9\%$ (mean $\pm$ s.e.m., $N = 8$), compared with untreated controls ($7 \pm 10\%$, $N = 8$). Cultures receiving dopamine applications, which can evoke short-lasting synaptic depression¹⁴, were not significantly different from controls ($5 \pm 13\%$, $N = 8$) when retested 24 h later (Fig. 1a, b). For all treatments, there was no significant change in the resting potential, action potential and input resistance measured in the motor cell L7 or in the capacity of the connection to be modulated further with application of transmitter or with homosynaptic stimulation (not shown).

Because long-term facilitation induced by repeated applications of 5-HT can be blocked by translational or transcriptional inhibitors present during the 5-HT exposures¹⁰, we next deter-

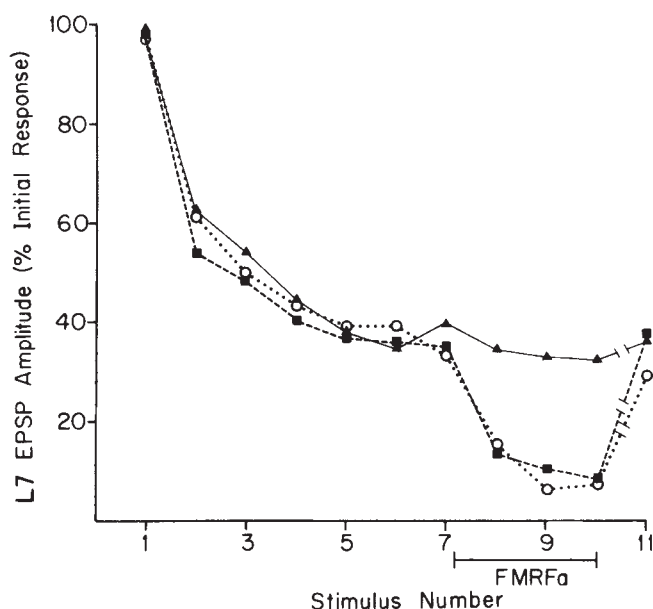


Fig. 3 Anisomycin does not interfere with short-term synaptic depression induced by either homosynaptic stimulation or FMRFamide application. Each point is the mean amplitude of the EPSPs normalized relative to the first evoked EPSP in each experiment. Sensory neurons were stimulated at 30-s intervals as described in Fig. 4b. In control cultures (closed triangles; $N=4$), the amplitude of the EPSPs reached a stable level at about 35–40% of the initial value after the sixth stimulus. A 2-min rest interval between the tenth and eleventh stimuli did not result in any appreciable recovery of the EPSP amplitude. Cells stimulated after treatment with anisomycin for 2 h (closed square; $N=4$) or 24 h after a 3-h exposure to anisomycin (open circles; $N=4$) undergo homosynaptic depression with kinetics similar to that of controls. Application of $0.1 \mu\text{M}$ FMRFamide when the amplitude of the EPSP had reached the stable level (either in the presence of anisomycin or 24 h after washout of anisomycin) reduced the EPSP amplitude to about 10% of the first evoked EPSP. The amplitude of the EPSP recovered to the steady-state level with a 2-min washout after the tenth stimulus.

mined whether the long-term depression of the EPSP in L7 produced by FMRFamide can be affected by blocking protein synthesis. Using the same FMRFamide application protocol, we exposed the cultures to anisomycin, a reversible translation inhibitor, at a concentration ($10 \mu\text{M}$) known to block protein synthesis by more than 95% (ref. 10). The presence of the inhibitor throughout the FMRFamide treatment blocked the decrease of the amplitude of the EPSP produced by FMRFamide to control levels when retested 24 h later; $12 \pm 12\%$ ($N=7$) compared with $14 \pm 10\%$ ($N=7$) for control cultures, and $-28 \pm 5\%$ ($N=7$) for FMRFamide-treated cultures (Fig. 2a and b). Anisomycin by itself, as was previously shown^{10,15}, had no significant effect on the EPSP or properties of the motor cells. In addition, anisomycin did not affect short-term synaptic depression. Brief application of FMRFamide either in the presence of anisomycin or 24 h after anisomycin washout, produced a short-lasting decrease in the amplitude of the EPSP (Fig. 3).

These results and the earlier findings by Montarolo *et al.*¹⁰ and Dale *et al.*¹³ on long-term synaptic facilitation indicate that the long-term efficacy of an identified synapse can be modified in opposite directions by 5-HT and FMRFamide. Moreover, long-term heterosynaptic modulation of this connection in either direction requires the synthesis of gene products not essential for short-term plasticity.

The opposing long-term actions of 5-HT and FMRFamide may result from their repeated activation of appropriate second-messenger systems used in the short-term process. Short-term

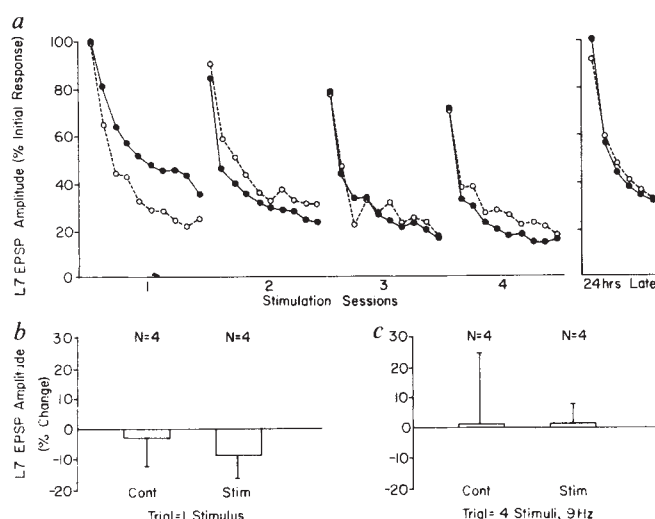


Fig. 4 Absence of long-term retention of synaptic depression following homosynaptic stimulation with two different protocols. *a*, A comparison of synaptic depression within and between stimulation sessions for experimental groups receiving one stimulus per trial (open squares) and four stimuli at 9 Hz per trial (closed square). Each point is the mean amplitude of the EPSPs of each experimental group ($N=4$) normalized as a per cent of the initial amplitude in each experiment. Repeated blocks of trials produced a comparable decrement in the EPSP amplitude. Following each block, there is a partial recovery of the response. Although there was significant depression of the EPSPs at the end of the last session, the amplitude of the EPSP was not significantly reduced when re-examined 24 h later. *b* and *c*, Summary of the per cent change in the amplitude of the EPSP after 24 h for control unstimulated cells and for cells stimulated with the single stimulus protocol (*b*) and the four stimuli at 9 Hz protocol (*c*). The height of each bar is the mean $\pm$ s.e.m. of the per cent change of the initial evoked EPSP compared with the first EPSP of the same connections re-examined 24 h later. Comparisons of the means for each experiment (two-tailed *t* test) indicated no significant difference between the control and experimental groups.

Methods. Each culture dish contained a single motor cell L7 and two mechanosensory cells. After 4–5 days, the amplitude of the EPSPs evoked by stimuli to each sensory neuron was measured as described in Fig. 1. One sensory neuron in each culture was stimulated only once on the first day (control) and re-examined 24 h later. The other sensory neuron in each culture was given four blocks of stimuli at 90-min intervals with each block consisting of ten trials at 30-s intervals. Each trial consisted of either a single stimulus sufficient to evoke a single action potential in the sensory cell ($N=4$) or four stimuli (four action potentials) at 9 Hz ($N=4$). After each block, the electrode was withdrawn from the L7 motor neuron. The experimental cells were retested 24 h later with six trials (one stimulus per trial) at 30-s intervals to determine both the amplitude of the EPSP and the capacity of the connection to undergo homosynaptic depression.

facilitation by 5-HT involves cyclic AMP-dependent phosphorylation^{16,17}, whereas short-term depression by FMRFamide involves the lipoxygenase metabolites of arachidonic acid⁷. Recent experiments using cell culture¹⁵ indicate that these two second messengers may also have a role in the long-term modulation of sensory-L7 connections. Treatment for ≈ 2 h with a cAMP analogue results in an enhancement of the amplitude of the EPSP in L7 lasting 24 h, whereas treatment with arachidonic acid for ≈ 2 h leads to a depression in synaptic strength lasting 24 h.

What is the behavioural role of this long-term synaptic depression? In the intact animal, repeated stimulation of the siphon skin with four blocks of 10 stimuli results in a marked habituation of reflex response 24 h later¹⁸ and is associated with a long-term depression of the connections between the sensory and motor neurons. Earlier work on the cellular analogues of

short-term habituation^{1,9} showed that one training session of 10 stimuli evoked in sensory neurons leads to a short-term depression of the connection between sensory and motor cells that parallels the short-term habituation. We have now examined whether repeated training sessions of stimuli in sensory neurons, using a protocol analogous to that used for long-term habituation, can, in turn, lead to long-term depression of the monosynaptic connection in cell culture (Fig. 4).

After we measured the amplitude of the EPSP evoked by each sensory neuron, one of the sensory cells in each culture was then stimulated at 90-min intervals with four blocks of 10 stimuli. Two stimulus protocols were used to reflect the range of effects evoked with moderate tactile stimulation to the siphon skin¹⁹. For both stimulation protocols, the build-up of the depression of the monosynaptic EPSP within each session and between the sessions (Fig. 4a) is comparable with that obtained in the intact nervous system for the synaptic input on to L7 with repeated stimulation of the siphon or the siphon nerve¹⁸. However, unlike the situation in the intact nervous system, both stimulation protocols failed to change significantly the amplitude of the EPSP when the same cells were retested 24 h later; $-8 \pm 7\%$ ($N = 4$) for the single stimulus protocol (Fig. 4b), and $1 \pm 8\%$ ($N = 4$) for the four stimuli protocol (Fig. 4c). These changes in the EPSP were not different than those measured for the unstimulated control cells ($-3 \pm 10\%$; $1 \pm 25\%$, respectively) in the same cultures (Figs 4b, c).

These *in vitro* results raise the possibility that the cellular mechanism underlying persistent synaptic depression observed in long-term habituation may involve heterosynaptic inhibitory influences such as those produced by modulatory transmitters as FMRFamide (see also refs 20, 21). This possibility parallels the finding in vertebrates, where short-term habituation can be produced by homosynaptic mechanisms, whereas long-term habituation cannot, but requires a central nervous system pathway not directly involved with the acquisition of the short-term forms of the behavioural modification^{22,23}. The recent identification of the neurons in *Aplysia* that produce presynaptic inhibition⁸ now permits a direct evaluation of the possible role of heterosynaptic inhibitory pathways in long-term habituation in the intact nervous system.

We wish to thank R. Woolley for technical assistance, A. Krawetz for typing the manuscript, the *Aplysia* Mariculture Facility of the Howard Hughes Medical Institute for providing all the animals used in this study, and Drs V. Castellucci, N. Dale, R. Hawkins, I. Kupfermann and Y. Yaari for their critical comments on the manuscript. This work was supported by the Howard Hughes Medical Institute (E.R.K.), the McKnight Foundation (S.S.), and the National Institutes of Health (S.S. and P.G.M.).

Received 2 February; accepted 6 April 1988.

- Castellucci, V. F., Pinsker, H., Kupfermann, I. & Kandel, E. R. *Science* **167**, 1745-1748 (1970).
- Castellucci, V. F. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **71**, 5004-5008 (1974).
- Castellucci, V. F., Carew, T. J. & Kandel, E. R. *Science* **202**, 1306-1308 (1978).
- Frost, W. N., Castellucci, V. F., Hawkins, R. D. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 9266-9270 (1985).
- Abrams, T. W., Castellucci, V. F., Camardo, J. S., Kandel, E. R. & Lloyd, P. E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7956-7960 (1984).
- Belardetti, F., Kandel, E. R. & Siegelbaum, S. A. *Nature* **325**, 153-156 (1987).
- Piomelli, D. *et al. Nature* **328**, 38-43 (1987).
- Mackey, S. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 8730-8734 (1987).
- Rayport, S. G. & Schacher, S. J. *Neurosci.* **6**, 759-763 (1986).
- Montarolo, P. G. *et al. Science* **234**, 1249-1254 (1986).
- Brunelli, M., Castellucci, V. F. & Kandel, E. R. *Science* **194**, 1178-1181 (1976).
- Mackey, S. L., Hawkins, R. D. & Kandel, E. R. *Soc. Neurosci. Abstr.* **12**, 1340 (1986).
- Dale, N., Schacher, S. & Kandel, E. R. *Science* **239**, 282-285 (1988).
- Montarolo, P. G., Kandel, E. R. & Schacher, S. *Soc. Neurosci. Abstr.* **13**, 391 (1987).
- Schacher, S., Castellucci, V. F. & Kandel, E. R. *Science* (in the press).
- Kandel, E. R. & Schwartz, J. H. *Science* **218**, 433-443 (1982).
- Ocorr, K. A., Tabata, M. & Byrne, J. *Brain Res.* **371**, 190-192 (1986).
- Carew, T. J. & Kandel, E. R. *Science* **182**, 1158-1160 (1973).
- Byrne, J., Castellucci, V. F., Carew, T. J. & Kandel, E. R. *J. Neurophysiol.* **41**, 402-417 (1978).
- Goldberg, J. I. & Lukowiak, K. *J. Neurobiol.* **15**, 395-411 (1984).
- Lockery, S. R. & Kristan, W. B. *Soc. Neurosci. Abstr.* **13**, 388 (1987).
- Leaton, R. N. & Supple, W. J. *Science* **232**, 513-515 (1986).
- Jordan, W. P. & Leaton, R. N. *J. comp. physiol. Psychol.* **96**, 170-183 (1982).

Facilitation of a dendritic calcium conductance by 5-hydroxytryptamine in the substantia nigra

S. Nedergaard*†, J. P. Bolam‡ & S. A. Greenfield*

* University Department of Pharmacology and ‡ MRC Anatomical Neuropharmacology Unit, South Parks Road, Oxford OX1 3QT, UK

Within the substantia nigra, the dendrites of dopaminergic neurons that project to the striatum appear to play an active and non-classical role in the physiology of the neuron in that they release transmitter and protein¹, but little is known of the factors controlling release of substances from these dendrites. In this study, we show that 5-hydroxytryptamine, which is contained in afferent fibres to the substantia nigra, is present in terminals making direct synaptic contact with dopaminergic neurons and also that it has a site-dependent, receptor-mediated, facilitatory effect on a specific dendritic calcium-dependent potential in nigrostriatal neurons *in vitro*. The ionic and spatial features of this response, which is insensitive to blockade by three different K⁺-channel antagonists, could correspond to those underlying the dendritic release of dopamine.

An important feature of nigrostriatal neurons is that dopamine (DA)² and acetylcholinesterase (AChE) (ref. 1) are released from the dendrites, but it is puzzling that this release is tetrodotoxin (TTX)-resistant² and unrelated to the overall excitability of the cell^{1,3}. In nigrostriatal neurons, a TTX-resistant calcium conductance is present which displays a high voltage threshold for generation from the cell soma (high-threshold calcium spike, HTSgCa²⁺) and which has been tentatively localized in the distal dendrites³. It is therefore possible that this conductance could mediate dendritic release of dopamine^{1,3}. In the substantia nigra, 5-hydroxytryptamine (5-HT) is present in afferent terminals⁴ and evokes the release of DA (ref. 2) and AChE (ref. 5). If the HTSgCa²⁺ is involved in the dendritic release mechanism, then administration of 5-HT in the vicinity of nigrostriatal cell dendrites should modify this conductance. The object of the present experiments was to test this hypothesis.

To determine whether an anatomical substrate exists for an interaction between 5-HT and dopaminergic neurons in the substantia nigra, sections of guinea pig mesencephalon were subjected to a double immunocytochemical procedure^{6,7} to localize both 5-HT and tyrosine hydroxylase (TH), one of the synthetic enzymes of catecholamines (Fig. 1). In the light microscope, the substantia nigra contained a fairly sparse distribution of 5-HT-immunoreactive axons, the density of which was somewhat higher in the ventral and medial aspects of the pars reticulata. The 5-HT-immunoreactive axons gave rise to swellings or boutons, which were seen in the electron microscope as vesicle-filled boutons forming asymmetrical synaptic specializations predominantly with proximal and distal dendrites and spines. Many of the immunoreactive boutons were juxtaposed to either dendrites or perikarya of neurons that were immunoreactive for TH and, when examined in complete sets of serial sections, were found to form synaptic specializations with them (Fig. 1).

During intracellular recordings from pars compacta neurons, (Fig. 2a), the HTSgCa²⁺ could be evoked by depolarizing current pulses to the cell soma at thresholds 10-40 mV positive to resting membrane potential (Fig. 2b) in the presence of TTX (10^{-6} M). The spike amplitude was proportional to the strength of the activating current pulse. When 10^{-2} M tetraethylammonium chloride (TEA) was added to the bathing solution, the HTSgCa²⁺ was evoked more easily (Fig. 2c), whereas a marked attenuation was produced by 10^{-3} M CdCl₂ (Fig. 2d). Applica-

† Permanent address: Institute of Physiology, University of Aarhus, DK-8000 Aarhus C, Denmark. Address correspondence to S.A.G.

Fig. 1 Electron micrographs of the substantia nigra pars reticulata of the guinea pig, incubated to reveal both tyrosine hydroxylase and 5-HT immunoreactivities by the peroxidase antiperoxidase technique, but using different chromogens for the peroxidase reactions^{6,7}. Structures containing 5-HT-immunoreactivity were first localized using 3,3'-diaminobenzidine as the chromogen for the immunoperoxidase reaction. Tyrosine hydroxylase-immunoreactive sites were then localized using benzidine dihydrochloride as the chromogen of the peroxidase reaction and the sections prepared for electron microscopy. Control reactions included single immunostaining for each antigen alone and the whole double-staining procedure, but with the omission of either or both of the primary antibodies. The reaction product formed by each chromogen is distinguishable in both the light and electron microscopes^{6,7}. *a*, A 5-HT-immunoreactive bouton (5-HT), which contains the typical amorphous diaminobenzidine reaction product, forms an asymmetrical synaptic contact (small arrows) with a small non-immunoreactive dendrite. In the lower half of the micrograph, a large dendrite contains the reaction product granules that are formed when benzidine dihydrochloride is used as a chromogen, thus identifying it as TH-immunoreactive (TH). *b* and *c*, 5-HT-immunoreactive boutons form asymmetrical synaptic contact (arrows) with TH-immunoreactive dendrites that contain the benzidine dihydrochloride reaction product granules. In *c* the reaction product has been partially lost during processing and only ghosts of the crystals remain. The immunoreactive dendrite in this section is also apposed to a bouton (*) that is not immunoreactive for 5-HT. Scales: *a* and *b* are at the same magnification: bar, 0.5 μm ; in *c*, bar is 0.2 μm .

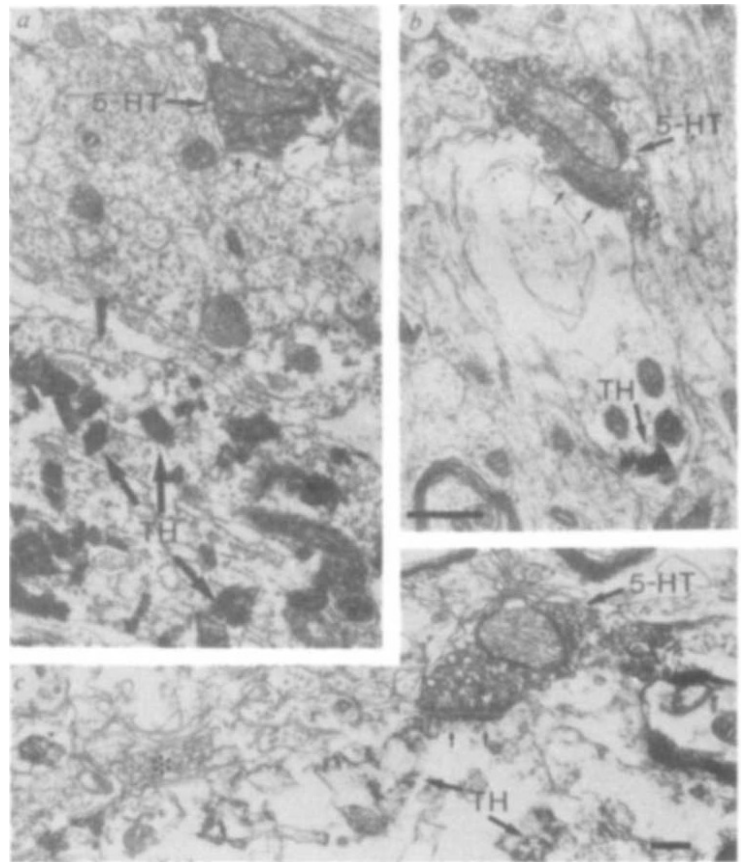
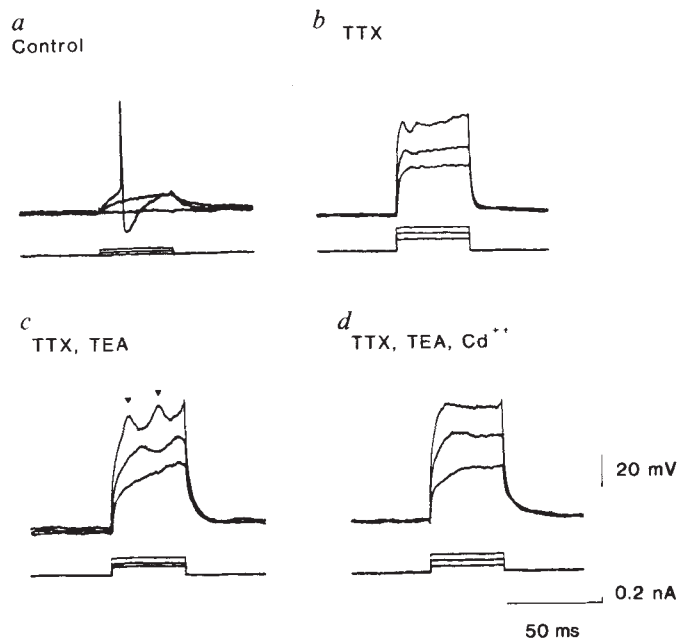


Fig. 2 Intracellular recordings from a single nigrostriatal neuron: depolarization from resting membrane potential (-57 mV) following blockade of various ion channels. *a*, Control: family of current pulses from a control period showing the voltage threshold for sodium-dependent action potential. Note the shape of the action potential is characteristic of a nigrostriatal neuron. *b*, TTX: in the presence of 10^{-6} M tetrodotoxin, strong currents evoke 'spikes' of short 5–10 ms duration (HTSgCa^{2+}), indicated by the pointer. Note the increased spike amplitude with increased current strength, and the significant after-hyperpolarization. *c*, TTX, TEA: in the presence of TTX and 10^{-2} M TEA, the membrane input resistance has increased. Smaller current pulses are now needed to evoke the HTSgCa^{2+} : their amplitude and duration is enhanced. *d*, TTX, TEA, Cd^{2+} : when 10^{-3} M Cd^{2+} is added with both TTX and TEA present, the HTSgCa^{2+} is markedly attenuated.

Methods. Brain slices were obtained from the guinea pig midbrain. Maintenance of slices and intracellular recordings were performed according to the methods described earlier³. Impaled neurons were identified as nigrostriatal cells by their high input resistance ($117 \pm 17\text{ M}\Omega$ (s.e.m.), $n = 14$), the presence of large after-hyperpolarizations, and the generation of a slow depolarizing conductance from membrane potentials negative to rest³. The resting membrane potentials were $-61.1 \pm 1.3\text{ mV}$ (s.e.m.), $n = 14$. TTX, TEA and Cd^{2+} were added directly to the perfusion medium.

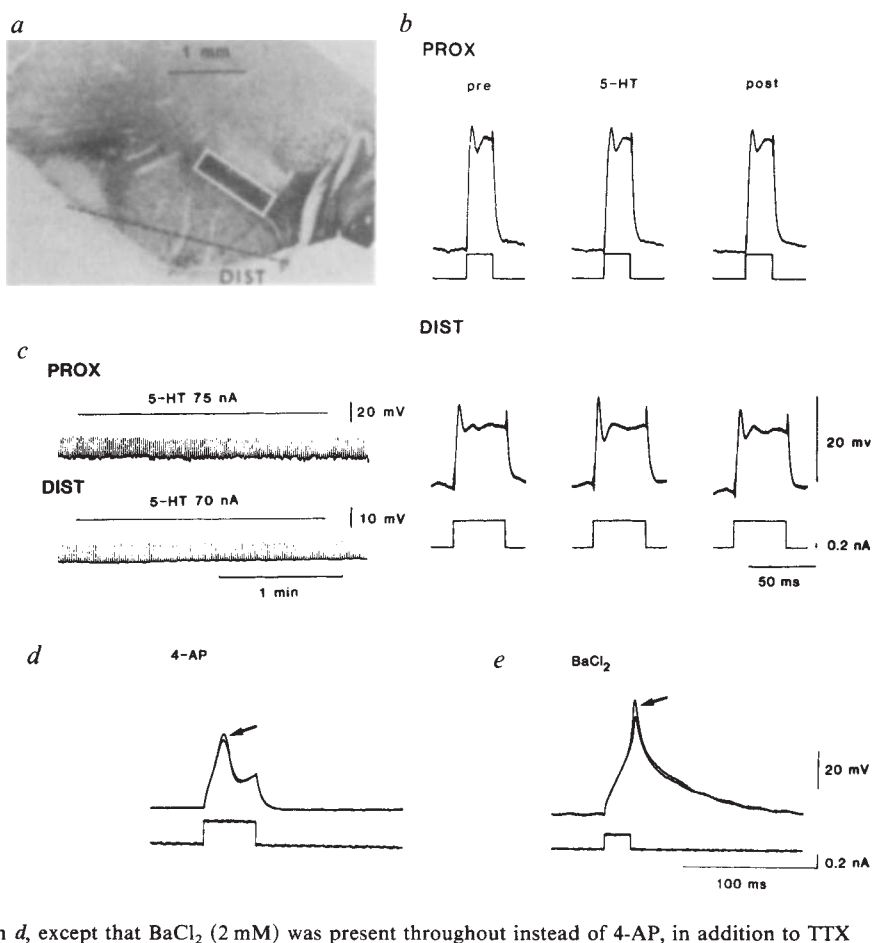


tion of BaCl_2 abolished the after-hyperpolarization, whereas 4-aminopyridine spared this after-hyperpolarization, yet increased the duration of the positive phase.

Iontophoretic application of 5-HT to the distal dendritic field, in the presence of TTX and TEA (Fig. 3*a*), produced a marked facilitation of the HTSgCa^{2+} in all the cells tested (Fig. 3*b*): in 33 applications to 14 cells, the average spike amplitude was increased to $124 \pm 4\%$ (s.e.m.). This response persisted in experiments where either BaCl_2 (spike amplitude increased to $125 \pm 5\%$ (s.e.m.), $n = 6$), or 4-AP ($120 \pm 4\%$ (s.e.m.), $n = 5$), or both,

were present in addition to TEA. Control ejections of Na^+ had no effect. But when 5-HT was applied iontophoretically into the region of the proximal dendrites (ejection current up to 100 nA; 4 applications to 3 cells), no significant effect on the HTSgCa^{2+} could be detected (Fig. 3*b*); the average spike amplitude was now $99 \pm 5\%$ (s.e.m.) of control values ($n = 4$). In four experiments where the 5-HT receptor antagonist⁸ cinanserin ($2.5 \times 10^{-5}\text{ M}$) was added to the perfusion medium, the 5-HT response in the distal dendrites was abolished. The membrane potential and input resistance of the cells did not change during ion-

Fig. 3 Effects of 5-HT application to localized portion of nigrostriatal cell dendrites. Three-barrel iontophoresis electrodes, containing 0.01 M 5-HT in two barrels ($pH=3.6$) and 3.0 M NaCl in one barrel, were used. TTX (10^{-6} M) and TEA (10^{-2} M) were present in all cases. *a*, Coronal section of guinea pig midbrain stained for TH, showing schematic area of recording (white box) and sites of application of 5-HT. Data also include recordings from more rostral sections. 5-HT was delivered by balanced applications (2–100 nA) to either of two different dendritic regions within the substantia nigra: (1) the proximal dendritic field of nigrostriatal cells at the pars compacta/pars reticulata borderline, 100–200 μm ventral to the recording site or (2) the distal dendritic field located in the ventral part of the pars reticulata, 500–1,000 μm from the recording site (DIST) ventral to the dotted line. *b*, Averaged membrane-potential responses to intracellular current pulses showing the HTSgCa $^{2+}$, before, during, and after a 75 nA application of 5-HT. In the proximal dendritic field (PROX), the HTSgCa $^{2+}$ remains unchanged. When 5-HT is applied (70 nA) in the distal dendritic field (DIST), the HTSgCa $^{2+}$ amplitude is increased by 61%. *c*, Chart records of cell membrane potential during repetitive depolarizing current pulses (TTX and TEA present). Iontophoretic applications of 5-HT to the proximal (PROX) and to the distal (DIST) dendritic fields does not lead in either case to changes in the resting membrane potential, (-60 mV and -54 mV respectively). *d*, Two superimposed averaged membrane potential responses in control conditions and during 5-HT application at 80 nA (arrowed). Throughout these recordings 4-AP was present in addition to TTX and TEA. *e*, As in *d*, except that BaCl $_2$ (2 mM) was present throughout instead of 4-AP, in addition to TTX and TEA.



Methods. 4-AP was added to the normal perfusing solution to a final concentration of 5 mM. In experiments where BaCl $_2$ was used, slices were first perfused with a control solution¹⁶ to eliminate precipitation. Subsequently this solution was changed to one where Ca $^{2+}$ ions were replaced with Ba $^{2+}$ (2 mM) (ref. 16). Throughout recordings, in the presence of TTX and TEA, HTSgCa $^{2+}$ were evoked by uniform depolarizing pulses at 1 Hz. Because the amplitude of two consecutive responses varied considerably, single responses were averaged in samples of 5, each obtained at 5-s intervals. The amplitude of the HTSgCa $^{2+}$ was measured on the averaged responses as the difference (in mV) between the peak of the spike and the maximal after-hyperpolarization. The mean amplitude of the HTSgCa $^{2+}$ was 12.8 ± 1.7 mV (s.e.m.), $n=14$. In the case of Ba $^{2+}$ -induced abolition of the after-hyperpolarization, amplitude was measured from peak of the spike to resting potential.

tophoretic applications of 5-HT to either the distal or proximal dendritic fields (Fig. 3c).

The onset of the effects of distal application of 5-HT occurred within a maximal latency period of 10 s, and a plateau phase developed after 120 s of this application (Fig. 4a). Individual cells showed an almost linear dose-response relation with moderate to strong ejection currents; weaker currents (2–20 nA) showed a bigger slope of the curve (Fig. 4b).

The HTSgCa $^{2+}$ is probably located in the distal dendrites of nigrostriatal cells, because the activation threshold is lowered under conditions where the electrical coupling between the dendrites and the soma is enhanced, that is during suppression of the potassium conductance by TEA (ref. 3) or 4-AP. Furthermore, the amplitude of the HTSgCa $^{2+}$ is not 'all or none', but proportional to current intensity, suggesting that it could be a summation of several smaller responses, each of which converge on the cell soma from different dendritic sites or branches³. Application of Cd $^{2+}$ attenuated the response, which can thus be attributed principally to calcium ions. In this study, measurements of HTSgCa $^{2+}$ amplitude included both positive and negative components of the transient. As reported for Na $^{+}$ -mediated somatic action potentials⁹, the after-hyperpolarization appears to be controlled by calcium entry, namely gK $^{+}$ Ca $^{2+}$, as it was abolished by BaCl $_2$ and could thus be readily differentiated from the initial inward current.

The increase in the HTSgCa $^{2+}$ amplitude is likely to be pro-

duced by an activation of cinanserin-sensitive 5-HT receptors, which are presumably localized at the site of synaptic specializations between the 5-HT containing boutons and the dopaminergic neurons. As this response was not abolished in the presence of three different K-antagonists, it is suggested that it might be independent of a gK $^{+}$. But the recorded Ca $^{2+}$ -dependent potentials, being markedly attenuated in somatic records, would be sensitive to even minor changes in local dendritic potassium or chloride conductances. It is therefore not possible to definitively exclude involvement of a residual dendritic K $^{+}$ or Cl $^{-}$ conductance. Indeed, 5-HT-induced suppression of a gK $^{+}$ has been reported in alysia¹⁰. Furthermore, evidence for a 5-HT induced modulation of a hyperpolarizing current analogous to the acetylcholine (ACh)-sensitive M-current¹¹ has been demonstrated in the hippocampus¹². On the other hand, a direct effect of 5-HT on the Ca $^{2+}$ -conductance underlying the HTSgCa $^{2+}$ could also explain the observed effect. Facilitation of a Ca $^{2+}$ -current by 5-HT has already been described in alysia neurons¹³.

Facilitation occurred only during applications to the distal, but not proximal, region of the dendrites: this observation implies that the conductances which control the HTSgCa $^{2+}$ and/or the activation of 5-HT receptors, occur in a localized and distant dendritic portion of these cells, despite the fact that 5-HT-immunoreactive boutons were found in contact with proximal as well as distal parts of the dendrite. It is possible that 5-HT has another action^{14,15}, which is either indirect via

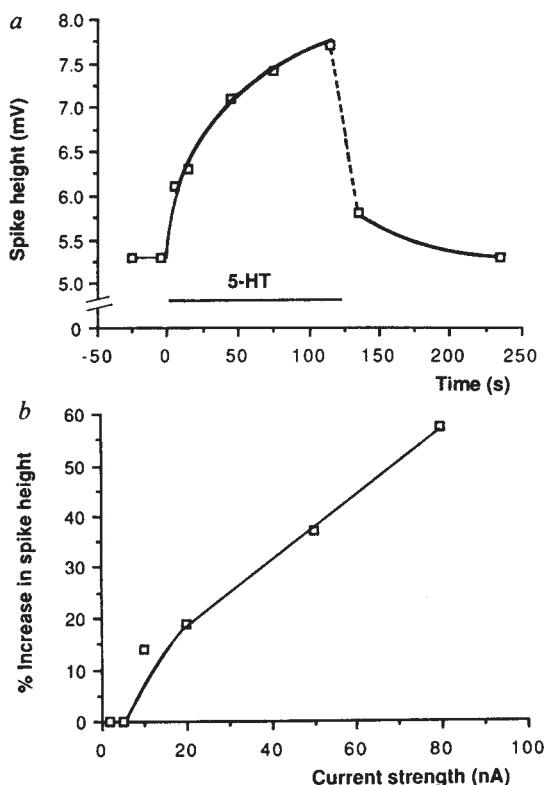


Fig. 4 *a*, Time course of the 5-HT effect. Average HTSgCa²⁺ amplitudes from different 10-s periods before, during, and after 5-HT (delivered with 80 nA for 120 s to the distal dendrites of a nigrostriatal cell). Note the rapid onset of the response (<10 s), and the course towards a plateau during prolonged application. *b*, Dose-response relation obtained from the same cell as in *a*. 5-HT was applied repeatedly with different current strengths from the same position in the distal dendritic field. Responses obtained after 120 s application were compared with those from control periods before and 120 s after 5-HT. Each point represents the normalized deviation, in HTSgCa²⁺ amplitude, from the control level (expressed as 100% for each application, irrespective of the actual value in mV).

result of contacts on the non-immunoreactive dendrites described in this study, or has an action in the more proximal regions which is not detected in the present experiments because of the presence of TEA.

In conclusion, the novel action of 5-HT reported here may have two fundamental implications: first, 5-HT can cause a significant change in a local dendritic Ca²⁺-dependent potential without provoking inhibition or excitation in the cell as a whole. Secondly, if this is a direct effect on the underlying Ca²⁺-current, it could represent the ionic mechanism by which 5-HT and other transmitters evoke the dendritic release of transmitter and protein¹ in the substantia nigra.

This work was supported by the Medical Research Council (UK), the Stanley Thomas Johnson Foundation, and E. P. Abraham Funds. S.N. was awarded a travel grant from MRC (Denmark). We would like to thank Chris Webb, Caroline Francis and Frank Kennedy for expert technical assistance. The Cinanserin used in this study was a gift from E.R. Squibb & Sons.

Received 7 March; accepted 14 April 1988.

- Greenfield, S. A. *Neurochem. Int.* **7**, 887-901 (1985).
- Glowinski, J. & Cheramy, A. in *Chemical Transmission: 75 Years* (eds Stjärne, L., Hedquist, P., Lagercrantz, H. & Wennmalm, A.) 127-149 (Academic, NY, 1981).
- Llinas, R., Greenfield, S. A. & Jahnsen, H. *Brain Res.* **294**, 127-132 (1984).
- Steinbusch, H. W. M. *Neuroscience* **6**, 557-618 (1981).
- Taylor, S. J., Bartlett, M. J. & Greenfield, S. A. *Neuropharmacology* **27**, 507-514 (1988).
- Levey, A. I. et al. *J. Histochem. Cytochem.* **34**, 1449-1457 (1986).
- Bolam, J. P. et al. *Brain Res.* **397**, 279-289 (1986).
- Roberts, M. T. T. & Straughan, D. W. *J. Physiol.* **193**, 269-294 (1967).

- Jahnsen, H. *Prog. Neurobiol.* **27**, 351-372 (1986).
- Siegelbaum, S. A., Camardo, J. S. & Kandel, E. R. *Nature* **299**, 413-417 (1982).
- Halliwel, J. V. & Adams, P. R. *Brain Res.* **250**, 71-92 (1982).
- Colino, A. & Halliwel, J. V. *Nature* **328**, 73-77 (1987).
- Pellmar, T. C. & Carpenter, D. O. *J. Neurophysiol.* **44**, 423-439 (1980).
- Dray, A., Gonye, T. J., Oakley, N. R. & Tanner, T. *Brain Res.* **113**, 45-57 (1976).
- Soubrie, P., Reisine, T. & Glowinski, J. *Neuroscience* **13**, 605-625 (1984).
- Llinas, R. & Sugimori, M. *J. Physiol.* **305**, 171-195 (1980).

Ankyrin and spectrin associate with voltage-dependent sodium channels in brain

Yogambal Srinivasan*, Lawrence Elmer*,
Jonathan Davis†, Vann Bennett†,
& Kimon Angelides*‡

* Department of Physiology and Molecular Biophysics,
Baylor College of Medicine, Houston, Texas 77030, USA

† The Howard Hughes Medical Institute and the Department of
Biochemistry, Duke University Medical Center, Durham,
North Carolina 27710, USA

The segregation of voltage-dependent sodium channels to specialized regions of the neuron is crucial for propagation of an action potential¹⁻⁵. Studies of their lateral mobility indicate that sodium channels are freely mobile on the neuronal cell body but are immobile at the axon hillock, presynaptic terminal and at focal points along the axon⁶. To elucidate the mechanisms that regulate sodium channel topography and mobility, we searched for specific proteins from the brain that associate with sodium channels. Here we show that sodium channels labelled with ³H-saxitoxin (STX) are precipitated in the presence of exogenous brain ankyrin by anti-ankyrin antibodies and that ¹²⁵I-labelled ankyrin binds with high affinity to sodium channels reconstituted into lipid vesicles. The cytoplasmic domain of the erythrocyte anion transporter competes for the latter interaction. Neither the neuronal GABA (γ-aminobutyric acid) receptor channel complex nor the dihydropyridine (DHP) receptor bind brain ankyrin. The results indicate that brain ankyrin links the voltage-dependent sodium channel to the underlying cytoskeleton and may help to maintain axolemmal membrane heterogeneity and control sodium channel mobility.

We purified sodium channels from rat brain membranes to an average specific activity of 2,200 ± 200 pmol ³H-STX binding mg per protein (Fig. 1) by sequential chromatography on DEAE-Sephadex, hydroxylapatite (HAP), and wheat germ agglutinin (WGA) Sepharose^{7,8}. Silver-stained gels run at various stages of the purification show enrichment of the diffusely staining α-peptide of relative molecular mass (*M_r*) ~220,000-260,000 (220-260K) with the absence of the 36K β1 or 33K β2 subunits^{8,9}. During the purification procedure we noticed the presence of two bands resembling the α- and β-subunits of brain spectrin of *M_r* 260 and 265K (ref. 10). The resolution and visibility of these polypeptides depended upon the gel system. In the Fairbank's 3.5-17.5% exponential gradient polyacrylamide gel (Fig. 1), these proteins are clearly resolved from the sodium-channel protein, whereas in the linear gradient system¹¹ routinely used to monitor sodium channel purity (ref. 7, Fig. 1), the spectrin bands migrate in the same position as the sodium channel and are masked by the diffuse staining of the sodium channel protein.

Although we tentatively ascribed the 260 and 265K bands to the α- and β-components of brain spectrin, we tested for both anti-spectrin and anti-ankyrin reactivity associated with the channel by immunoblotting because ankyrin is often associated with spectrin in erythrocyte membranes. Figure 1 shows that both spectrin and the 220K form of rat brain ankyrin¹² are

‡ To whom correspondence should be addressed.

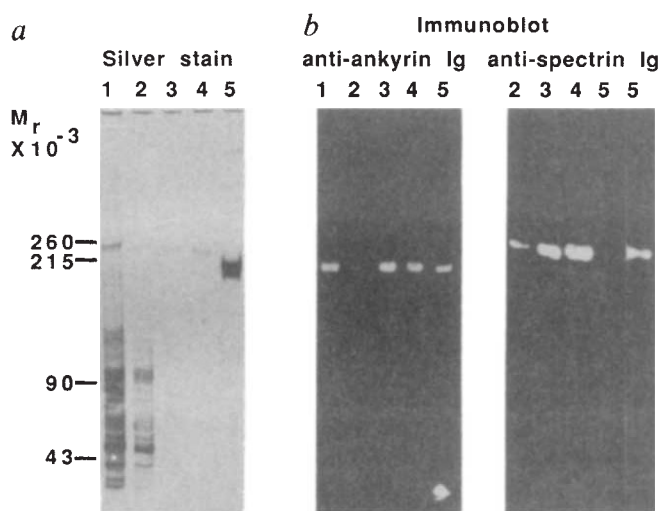


Fig. 1 Analytical SDS-PAGE of rat brain sodium channel purification (a) and immunoblots of spectrin and ankyrin in this preparation (b). Lanes loaded with 5 μ g protein. Lane 1, rat brain membranes showing spectrin doublet at 260–265K; lane 2, Triton X-100 extract; lane 3, peak fraction from DEAE-Sephadex; lane 4, HAP peak fraction; lane 5, WGA-Sephadex. Samples electrophoresed in three sets of Fairbank's 3.5–17.5% exponential gradient polyacrylamide gels. The immunoblots were probed with anti-ankyrin or anti-spectrin antibodies at dilutions of 1:1,000. In the anti-spectrin immunoblot, the second lane (denoted 5) is an overexposure of lane 5. Control blots probed with non-immune IgG do not show any reactivity. Ankyrin and brain spectrin antibodies were raised and characterized as described¹⁵. The minor band in lane 1 is the 190K breakdown product of brain ankyrin^{12,15}. Spectrin and ankyrin were consistently found in each of the 12 preparations of sodium channels tested.

Methods. The sodium channel was purified from rat brain by sequential chromatography on DEAE-Sephadex, hydroxylapatite (HAP) and wheat-germ agglutinin (WGA)-Sephadex columns as described elsewhere^{7,8}. ³H-STX was used to assay for the sodium channel by the rapid centrifugation on Sephadex G-50 columns²⁴. Between 100–200 μ g protein was typically obtained from the WGA step and the average specific activity of six preparations is $2,200 \pm 100$ pmol ³H-STX binding per mg protein. Protein concentrations were determined by the method of Peterson²³ and silver staining was by the method of Merrill²⁵. For immunoblotting, the proteins were transferred from polyacrylamide gels to nitrocellulose and the blots blocked with 4% bovine serum albumin in 10 mM-sodium phosphate, 0.15 M NaCl, 1 mM EDTA, 0.2% Triton X-100, pH 7.5 (PBS) for 1 h and incubated with 1:1,000 final dilution of affinity-purified anti-ankyrin or anti-spectrin, or with non-immune IgG overnight at 4°C. Blots were washed with PBS three times for 10 min each and once with 2 M urea, 0.1 M glycine, 1% Triton X-100, once more with PBS and incubated with 2 ml ¹²⁵I-Protein A (4×10^6 c.p.m.) for 2 h at 4°C in PBS. The blots were washed again using the same procedure as above and autoradiographed using pre-flashed Kodak XRP film.

present throughout the purification of the sodium channel protein. By densitometry, the approximate ratio of ankyrin:sodium channel is 0.9:1.0 in the HAP eluent and is reduced to 0.1:1.0 in the material eluting from the WGA step. Under the conditions used to wash the WGA-Sephadex column, 90% of the ankyrin and >99% of the spectrin is removed (Fig. 1, lane 5). The persistence of both spectrin and ankyrin in sodium channel preparations despite HAP and WGA chromatography suggests that these proteins could be associated with sodium channels.

To evaluate how ankyrin or spectrin bind to the sodium channel protein, we labelled highly enriched preparations of solubilized sodium channels with ³H-STX, a specific ligand for this ion channel, and tried to immunoprecipitate this activity with anti-ankyrin antibodies. We incubated ³H-STX-saturated sodium channels with increasing amounts of brain ankyrin plus

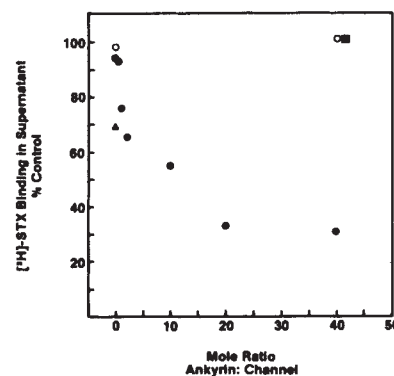


Fig. 2 Association of ankyrin with the sodium channel. Depletion of ³H-STX binding activity from supernatants of purified sodium channel in increasing concentrations of added ankyrin (●) by anti-ankyrin antibodies and by non-immune IgG (○, controls). (▲), Partially pure sodium channel sample from the HAP column-depletion of ³H-STX binding activity by anti-ankyrin antibodies without added ankyrin; Ankyrin plus anti-ankyrin antibodies do not deplete (■) ³H-flunitrazepam labelled GABA or (□) ³H-PN 200 labelled DHP receptors.

Methods. Purified sodium channel, 10 μ g, from the peak fraction of the WGA-Sephadex column was incubated with ³H-STX (22 nM) with and without 1.5 μ M tetrodotoxin for 20 min. in 200 μ l 100 mM choline chloride, 5 mM NaCl, 10 mM CaCl₂, 20 mM HEPES/Tris, 0.1% Triton X-100, 0.025% phosphatidylcholine, pH 7.4, and the freshly prepared protease mixture. Varying amounts of brain ankyrin in 10 mM sodium phosphate, 1 mM EGTA, 0.05% Tween, 1 mM Na₂S₂O₈, 10% sucrose, pH 7.4 was added to samples and incubated for 2 h at 4°C with shaking. Immune or non-immune IgG (1:1,000 final dilution) was added and the incubation continued at 4°C for 10 h before addition of 50 μ l of a 1:1 slurry of Protein A-Sepharose in binding buffer to each sample. After incubation for 2 h at 4°C samples were centrifuged for 2 min at 14,000 r.p.m. The supernatant was assayed for ³H-STX bound to sodium channels by rapid centrifugation on G-50 columns as described²⁴. ³H-STX binding was measured in control samples not treated with ankyrin or anti-ankyrin IgG, or treated with non-immune IgG and Protein A-Sepharose only, both at the beginning and at the conclusion of the experiment. The activity of the channel remained unchanged during this period. The average specific activity of the two preparations used in this presentation was $2,250 \pm 206$ pmol mg⁻¹ protein. Non-specific binding was always <1–2% of specific binding. Although the absolute level of ³H-STX binding that was precipitated varied slightly, the average value from six preparations was $60 \pm 15\%$.

saturating anti-ankyrin or non-immune IgG, and precipitated immune complexes with Protein A-Sepharose. Although some (typically 5–15%) of the ³H-STX binding activity can be depleted with anti-ankyrin antibodies without added ankyrin, adding ankyrin causes a concentration-dependent depletion of the sodium-channel activity in the supernatant (Fig. 2). Some of the ³H-STX binding activity can also be depleted with anti-ankyrin antibodies but without added ankyrin from peak fractions from the HAP column (Fig. 2, ▲) supporting the immunoblot observation that more ankyrin is present at this intermediate stage of the purification.

This interaction between ankyrin and the sodium channel is specific, because addition of other cytoskeletal proteins in place of ankyrin (such as neurofilaments or actin) does not deplete ³H-STX binding activity. Titration of ankyrin with purified sodium channels in solution indicates that saturation of ankyrin binding is achieved at a 40 mole excess of ankyrin over sodium channel with a $K_{0.5}$ of 100 nM under these conditions. Given the complex equilibria involved, the possibility that the antibody blocks some of the binding sites on ankyrin, and the possibility of partial denaturation of the sodium channel protein due to ankyrin binding, the fact that as much as 70% of the ³H-STX binding activity could be removed from solution by anti-ankyrin

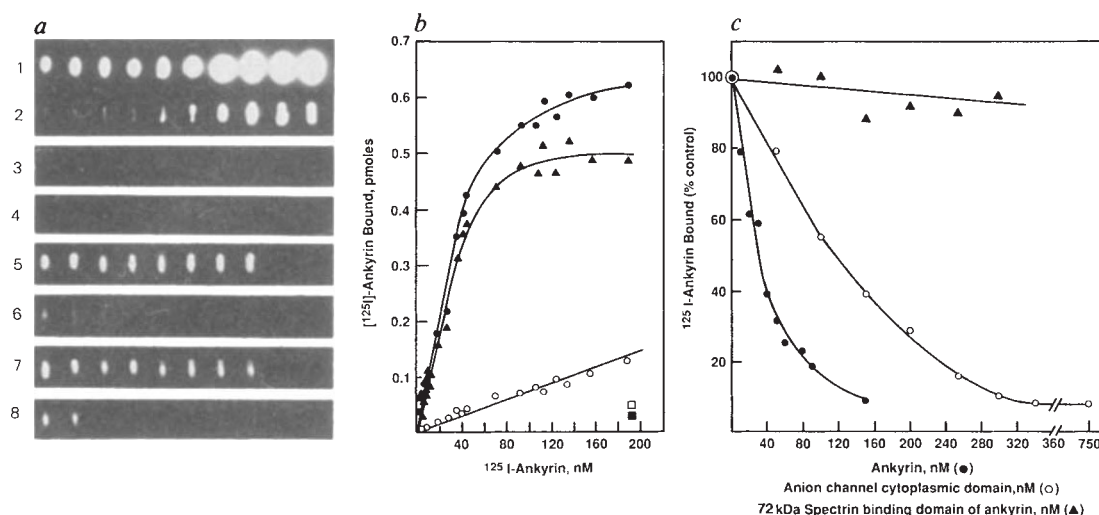


Fig. 3 Binding of ^{125}I -labelled ankyrin to the reconstituted sodium channel. Purified sodium channels were reconstituted into phosphatidylcholine/phosphatidylethanolamine (10:1) vesicles⁷ which were then absorbed onto nitrocellulose paper (0.4 μg channel protein). Varying concentrations of ^{125}I -labelled ankyrin were added to wells of a dot blot apparatus in the absence or presence of 250 nM unlabelled ankyrin. Parallel controls were performed with vesicles that contained no protein. The nitrocellulose blots were autoradiographed, and the paper in the wells then excised and counted in a gamma counter. *a*, Autoradiogram of ^{125}I -ankyrin binding to the sodium channel. Lane 1, increasing concentrations of ^{125}I -ankyrin (total binding); lane 2, binding in 250 nM unlabelled ankyrin (non-specific binding); lane 3, the binding to vesicles without protein; lane 4, the binding to reconstituted sodium channels whose cytoplasmic domain has been removed after digestion with 0.02% trypsin. Lanes 5-8 show displacement of 10 nM ^{125}I -ankyrin (lane 5), with increasing concentrations of unlabelled ankyrin (lane 6), the 72K spectrin-binding fragment of ankyrin (lane 7), and the 43K cytoplasmic fragment of band 3 (lane 8). *b*, Quantitative data from the autoradiograms in *a*. Total (●), non-specific (○) and specific (▲) binding of ^{125}I -labelled ankyrin to the sodium channel is shown. Also shown is the binding to vesicles without protein (■) and to reconstituted channels with the cytoplasmic domain removed (□). (*c*), Quantitative data on displacement from *a*, lanes 5-8 showing competition by unlabelled ankyrin (●), the 72K spectrin binding domain of ankyrin (▲), and the 43K cytoplasmic fragment of Band 3 (○).

Method. To make sodium channels lacking the cytoplasmic domain, trypsin digestion was carried out for 30 min then terminated with 0.1% soybean trypsin inhibitor and the digested fragment was removed by centrifuging the protein-containing vesicles at 100,000g for 30 min. As ^3H -STX binds to the extracellular face of the sodium channel, we verified that >40% of the ^3H -STX binding was retained after digestion and removal of the cytoplasmic domain by resolubilizing the vesicles in 0.2% Triton X-100, 0.05% phosphatidylcholine and carrying out the rapid centrifugation assay on Sephadex G-50 columns.

antibodies is remarkable. In addition to the reasons cited above, it is possible that the 30% of the ^3H -STX binding sites that are not immunoprecipitated are associated with sodium channel isoforms that do not bind brain ankyrin. We could also precipitate $\approx 20\%$ of the ^3H -STX binding activity from the supernatant using brain or erythrocyte spectrin that was covalently attached to Sepharose 6 MB beads.

Although it is known that brain ankyrin binds tightly to neuronal membranes^{13,14}, it does not interact with all membrane channels. Identical experiments performed with the rat brain benzodiazepine GABA receptor complex labelled with ^3H -flunitrazepam and with DHP receptors labelled with ^3H -PN 200 show no precipitation of these complexes with saturating ankyrin concentrations and anti-ankyrin antibodies (Fig. 2).

Depletion of ^3H -STX binding by ankyrin plus anti-ankyrin antibodies provides additional but indirect evidence that ankyrin and spectrin associate with the sodium channel. We therefore measured the direct binding of ^{125}I -labelled ankyrin to purified sodium channels reconstituted into lipid vesicles. These were immobilized on nitrocellulose paper and titrated with increasing concentrations of ^{125}I -labelled ankyrin. Non-specific binding was measured by titrating in the presence of saturating amounts of unlabelled ankyrin. Figure 3*a* shows that ^{125}I -ankyrin binds saturably to the sodium channel protein (lane 1), that non-specific binding is very low (lane 2), and that there is no binding of ankyrin to vesicles without protein (lane 3) or to the reconstituted channel in which the exposed cytoplasmic portion of the channel has been removed by digestion with trypsin (lane 4). Scatchard analysis of the direct binding data indicate a binding site with a K_d of 20 nM and a maximal binding capacity of 0.53 pmol under these experimental conditions. As the channel is reconstituted in a random orientation, where half the

channels have exposed cytoplasmic surfaces (determined by ^3H -STX binding), this indicates a stoichiometry of 1.1 moles ankyrin per mole of sodium channel. The binding between ^{125}I -ankyrin and the reconstituted sodium channel can be displaced by unlabelled ankyrin (Fig. 3*a*, lane 6; Fig. 3*c*, closed circles), or by the 43K cytoplasmic fragment of the erythrocyte anion transporter band 3 (which contains the ankyrin binding site) (Fig. 3*a*, lane 8; Fig. 3*c*, open circles) with a K_d of 14 nM and a $K_{0.5}$ of 150 nM. Although K_d s calculated from direct binding of ^{125}I -ankyrin or from competition with unlabelled ankyrin to the sodium channel are in good agreement, the K_d determined by displacement is slightly lower and may indicate that a small fraction of the ^{125}I -labelled molecules are inactive. Additional control experiments show that there is no displacement of ^{125}I -ankyrin binding to the sodium channel by the 72K spectrin binding domain of ankyrin¹⁵ (Fig. 3*a*, lane 7; Fig. 3*c*, triangles), indicating that the binding of ^{125}I -ankyrin is not to spectrin that remains with the sodium channel. Taken together, these data provide evidence that ankyrin binds to and associates with the voltage-dependent sodium channel and that this binding is of high affinity, stoichiometric, and through the cytoplasmic domain of the sodium channel. Furthermore, the sodium channel binding site shares some properties with the 43K cytoplasmic domain of band 3.

The work described here is the first report of the association of ankyrin with a neuronal membrane protein. Recently, binding of ankyrin to the kidney ($\text{Na}^+ - \text{K}^+$)ATPase, a membrane protein which assumes a non-uniform distribution in epithelial cells, has been reported¹⁶. Although direct binding of ^{125}I -labelled brain ankyrin or erythrocyte ankyrin to ankyrin-depleted neuronal membranes has shown specific membrane sites for this molecule¹²⁻¹⁴, the ankyrin-binding proteins were not identified.

Rotational diffusion measurements show that 70% of the sodium channels may be linked to the intermediate anchoring cytoskeleton, and fluorescence microscopy and photobleach recovery⁶ show that in specific regions of the neuron, voltage-dependent sodium channels are clustered and immobilized. Ankyrin is likely to be involved in immobilization of sodium channels because it is a multifunctional protein containing binding sites for tubulin, spectrin and membrane proteins¹⁷. In native membranes, there are at least 10 times more ankyrin molecules than sodium channels¹⁷ and the binding of the purified components is nearly stoichiometric. The relative abundance of ankyrin in cells probably reflects its attachment to other ion channels, receptors or cell-adhesion molecules.

The molecular composition of the sodium channel has been the subject of continuing controversy⁸. Only the α -peptide is required for its ion conduction properties¹⁸ and its ability to bind ankyrin. Interestingly, although the functional domains of the sodium channel and the DHP-binding protein are highly homologous, the regions putatively ascribed to cytoplasmic domains are not conserved¹⁹. This could partly explain the negligible interaction between the DHP-binding protein and brain ankyrin.

There is molecular specificity in defining the topography of ion channels over the nerve cell surface and in regulating their mobility. Double-label fluorescence microscopy and photobleach recovery studies of benzodiazepine/GABA receptors and sodium channels on spinal-cord neurons show that benzodiazepine/GABA receptors are segregated and immobilized on the neuronal cell body and the dendritic processes²⁰, whereas sodium channels are freely mobile in the cell body but are immobilized at the axonal hillock and presynaptic terminal. Furthermore, we show here that, at least for the GABA and DHP receptors, which both exhibit different cellular distributions and lateral mobilities, this form of brain ankyrin does not interact with the solubilized complexes.

Linkage of the sodium channel to ankyrin and subsequently to spectrin has the potential to either stabilize channels following assembly or to participate in the initial placement of channels at nodes of Ranvier. Such a role for spectrin or ankyrin in maintaining specialized distribution of integral membrane proteins has not been emphasized previously as in some cells spectrin or ankyrin molecules have been observed in an apparently uniform pattern under the plasma membrane¹³. In polarized cells such as transporting epithelia, photoreceptors and neurons, however, the distribution of ankyrin analogues is not uniform^{21,22}. In neurons, ankyrin is restricted to the membranes of the cell body and axons, and is not detected by immunostaining along dendrites where voltage-dependent sodium channels are sparsely distributed. Several mechanisms for heterogeneity within this lawn of spectrin and ankyrin have been described recently and indicate that spectrin and its associated proteins could, in principle, regulate the placement and restrict the lateral mobility of a variety of integral membrane proteins. Brain tissue contains at least two distinct isoforms of spectrin²⁶ and ankyrin²⁷ with different localizations and different programmes of expression during differentiation. Post-translational modifications of spectrin and ankyrin include phosphorylation, acylation with fatty acids, and limited proteolysis^{28,29} and could provide an additional basis for molecular sorting among potential integral proteins. It will be important in future experiments to characterize the population of ankyrin associated with the sodium channel, their localizations in nerve cells, and to determine if either the ankyrin or sodium channel has been modified or are specialized isoforms.

Supported by grants from the NIH and the Muscular Dystrophy Association.

3. Kristol, C., Sandie, C. & Akert, K. *Brain Res.* **142**, 391–400 (1978).
4. Rosenbluth, J. J. *Neurocytol.* **5**, 731–745 (1976).
5. Chiu, S. Y. & Ritchie, J. M. *Adv. Neurol.* **31**, 313–328 (1981).
6. Angelides, K. J., Elmer, L. W., Loftus, D. & Elson, E. L. *J. Cell Biol.* (in the press).
7. Elmer, L. W., O'Brien, B. J., Nutter, T. J. & Angelides, K. J. *Biochemistry* **24**, 8128–8137 (1985).
8. Hartshorne, R. P. & Catterall, W. A. *J. biol. Chem.* **259**, 1667–1675 (1984).
9. Catterall, W. A. *Science* **223**, 653–661 (1984).
10. Davis, J. Q. & Bennett, V. *J. biol. Chem.* **258**, 7757–7766 (1983).
11. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
12. Davis, J. Q. & Bennett, V. *J. biol. Chem.* **259**, 16198–16206 (1986).
13. Bennett, V., Baines, A. J. & Davis, J. Q. *Cell Biochem.* **29**, 157–169 (1985).
14. Davis, J. & Bennett, V. *J. biol. Chem.* **261**, 16198–16206 (1986).
15. Davis, J. & Bennett, V. *J. biol. Chem.* **259**, 1874–1881 (1984).
16. Nelson, W. J. & Veshnock, P. J. *Nature* **328**, 533–536 (1987).
17. Davis, J. & Bennett, V. *J. biol. Chem.* **259**, 13550–13559 (1984).
18. Stumer, W., Methfessel, C., Sakmann, B., Noda, M. & Numa, S. *Eur. Biophys. J.* **14**, 131–138 (1987).
19. Tanabe, T. *et al. Nature* **328**, 313–318 (1987).
20. Angelides, K. J., Velasquez, J., Thompson, C. & Barnes, E. M. *Adv. biochem. Pharmac.* (in the press).
21. Drenckhahn, D. & Bennett, V. *Eur. J. Cell Biol.* **43**, 479–486 (1987).
22. Nelson, W. J. & Veshnock, P. J. *J. Cell Biol.* **104**, 1527–1537 (1987).
23. Peterson, G. L. *Analyt. Biochem.* **83**, 346–356 (1977).
24. Levinson, S. R., Curatalo, C. J., Reed, J. K. & Raftery, M. *Analyt. Biochem.* **99**, 72–76 (1979).
25. Merrill, C. R., Dunau, M. L. & Goldman, D., *Analyt. Biochem.* **110**, 201–207 (1981).
26. Reiderer, B. M., Zagon, I. S. & Goodman, S. R. *J. Cell Biol.* **102**, 2088–2097 (1986).
27. Nelson, W. J. & Lazarides, E. *Cell* **39**, 309–320 (1984).
28. Staufenbiel & Lazarides, E. *Proc. natn. Acad. Sci., U.S.A.* **83**, 318–322 (1986).
29. Hall, T. & Bennett, V. *J. biol. Chem.* **262**, 10537–10545 (1987).

Development of CD4⁺CD8⁺ cytotoxic T cells requires interactions with class I MHC determinants

Suzana Marušić-Galešić*, David A. Stephany†, Dan L. Longo* & Ada M. Kruisbeek*

* Biological Response Modifiers Program, National Cancer Institute, and † Biological Resources Branch, National Institute for Allergy and Infectious Disease, National Institute of Health, Bethesda, Maryland 20892, USA

Differentiation of bone marrow derived precursors into mature T cells takes place in the thymus. During differentiation, T cells develop the receptor repertoire which allows them to recognize antigen in the context of self major histocompatibility complex (MHC) molecules. Mature T helper cells (mostly CD4⁺CD8[−]) recognize antigen in the context of class II MHC molecules, whereas cytotoxic T cells (mostly CD4[−]CD8⁺) recognize antigen in the context of class I MHC determinants^{1,2}. Thymic MHC-encoded determinants greatly influence the selection of the T-cell receptor repertoire^{3–6}. In addition to positive selection, a negative selection to eliminate self-reactive T-cell clones is thought to occur in the thymus⁷, but how this 'education' occurs is not well understood. It has been suggested that during differentiation an interaction between the T-cell receptor (TCR) and MHC-encoded determinants occurs, leading to the selection of an MHC-restricted receptor repertoire^{3–5}. In support of this hypothesis, class-II-specific, CD4⁺CD8[−] helper T cells fail to develop in mice neonatally treated with anti-class II monoclonal antibody (mAb)^{8,9}. As CD4[−]CD8⁺ cells differ from the CD4⁺CD8[−] lineage (in function, MHC-restriction specificity and perhaps site of education^{1,2,10}) we examined whether interactions with MHC determinants are also necessary for the development of class-I-specific T cells. Here we show that mice chronically treated with anti-class I mAb from birth lack CD4[−]CD8⁺ cells and cytotoxic T-cell precursors, indicating that most CD4[−]CD8⁺ T cells need interaction with class I MHC molecules during differentiation.

We treated C3H mice (H-2^k) from birth with anti-class I mAb¹¹ in daily doses sufficient to saturate the class I molecules on spleen and thymus cells of adult mice, as judged by flow cytometry (FCM, data not shown). Two- to three-week-old anti-class I treated mice appeared as healthy as saline-injected control animals. Two-colour FCM analysis of normal spleen cells using anti-CD4 and anti-CD8 mAb revealed two T-cell

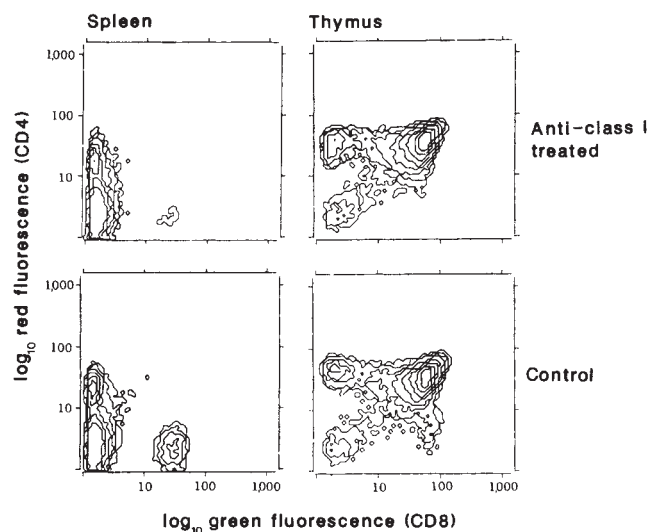


Fig. 1 Two-colour flow cytometry (FCM) analysis of cell surface CD4 and CD8 expression on spleen cells (left) and thymocytes (right) from anti-class I treated (top) and control (bottom) C3H mice. Cells were stained first with FITC conjugated anti-CD8 mAb (Becton-Dickinson), then washed and stained with biotin-conjugated anti-CD4 mAb², followed by Texas Red streptavidin (TRA) (BRL).

Methods. Neonatal C3H (H-2^k) mice were treated within 24 h of birth with anti-H-2K^d mAb (hybridoma 15-1-5, ref. 11, obtained from ATCC purified from ascites by ammonium sulphate precipitation followed by size separation. Mice were injected daily intraperitoneally with 500 µg of mAb per body weight and were analysed at 17–24 d in six separate experiments. Control mice were treated daily with saline, after preliminary experiments demonstrated that daily injection with equally high amounts of non-binding mouse anti-class I mAb (34-1-2, ref. 14) had no effect (data not shown). FCM analysis was as described¹², using a B-D Dual Laser FACS II interfaced to a PDP 11/24 computer. Data were collected on 50,000 viable cells (as determined by forward light scatter and propidium iodide staining) and displayed as contour diagrams, with a three-decade log scale of increasing green fluorescence on the x-axis and red fluorescence on the y-axis. The x and y coordinates which defined negative cells were selected as the intersection of positive and control negative profiles in each parameter. Negative controls were obtained by staining with an irrelevant control mAb. Per cent CD4⁺CD8⁺ cells: for control spleen, 5.7; for anti-class I spleen, 5.2; for control thymus, 11.1; for anti-class I thymus, 9.71; per cent CD4⁺CD8⁺ cells: for control spleen, 2.16; anti-class I spleen, 0.13; control thymus, 1.78; anti-class I thymus, 0.61.

subsets, CD4⁺CD8⁺ and CD4⁺CD8[−] (Fig. 1), as well as many CD4[−]CD8[−] cells (mostly B cells, data not shown). The number of CD4⁺CD8⁺ T cells was reduced in six independent experiments, by 4- to 10-fold in the spleen of anti-class I treated mice as compared with normal, age-matched controls, whereas the number of CD4⁺CD8[−] T cells was normal (see Fig. 1 legend). In the thymus, where T cells mature, four populations of T cells can be detected in normal mice: CD4⁺CD8[−], early T-cell precursors; CD4⁺CD8⁺, small cortical lymphocytes; CD4⁺CD8⁺, mostly mature cytotoxic T cells; and CD4⁺CD8[−], mature helper T cells¹². In anti-class I treated mice the CD4⁺CD8⁺ T cells, but not the other three types, were reduced in number by 3- to 4-fold (Fig. 1). Thus, development of CD4⁺CD8⁺ T cells in spleen and thymus is impaired in anti-class I treated mice, but development of CD4⁺CD8[−] and CD4[−]CD8[−] cells is not affected.

These results suggested that blocking of class I molecules during early post-natal development might prevent generation of mature CD4⁺CD8⁺ T cells. Alternatively, this treatment could somehow result in downregulation of the CD8 marker on this subset of T cells, without affecting function. As most alloreactive

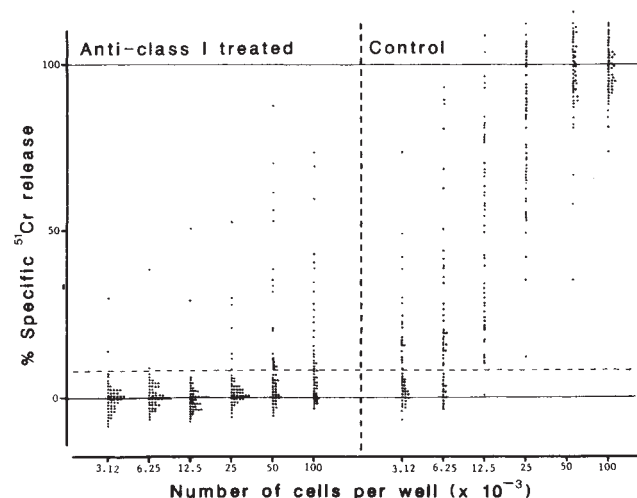


Fig. 2 Limiting-dilution analysis of CTL-precursor frequency in the spleen of anti-class I treated (left) and control (right) C3H mice. Groups of 48 replicate microcultures of varying numbers of C3H responder spleen cells (H-2^k) were cultured in the presence of 5×10^5 irradiated (2,000 R) BALB/c stimulator spleen cells (H-2^d) and 20% delectinated rat Con-A supernatant (rat T-cell monoclonal, Collaborative Research). On day 7, 10^3 ⁵¹Cr-labelled P815 (H-2^d) target cells were added per well, and ⁵¹Cr release in the supernatant was measured after 6 h. Each dot represents % specific ⁵¹Cr-release of an individual microwell. Positive wells were those in which ⁵¹Cr-release exceeded by >3 s.d. (broken line) the mean of control wells in which P815 cells were incubated with stimulator cells alone (solid line). Data are representative of four experiments. Precursor frequency values were calculated as described^{13,21} and are presented in Table 1.

Methods. Neonatal C3H mice were treated as described in Fig. 1, and analysed at 20 d. Cultures were set up in 96-well microtitre plates in a final volume of 0.2 ml of Eagle's Hanks amino-acid medium with 10% FCS 2 mM L-glutamine and 5×10^{-5} M 2-mercaptoethanol. Release of ⁵¹Cr was measured after centrifugation of the plates; 0.15 ml of supernatant was removed from each well and radioactivity of the samples was measured in a gamma counter.

cytotoxic T cells have a CD4[−]CD8⁺ phenotype, we measured the cytotoxic T lymphocyte (CTL) response of control and treated mice in bulk mixed lymphocyte cultures (MLC) against MHC alloantigens. Surprisingly, after five days of culture with allogeneic stimulator cells, we observed normal cytotoxicity with responder spleen cells from treated mice (data not shown) and FCM analysis revealed that this cell population contained equal numbers of CD4⁺CD8⁺ T cells in both control and anti-class I treated mice. There are two possible explanations: either the CD8 molecule is downregulated in treated mice but reexpressed during culture, or a few remaining CD4⁺CD8⁺ cells expand during the culture period. To distinguish these possibilities, the ability of anti-class I treated mice to generate a CTL-response to allo-class I antigens was measured in a sensitive limiting dilution culture system¹³. The number of CTL precursors in treated animals was reduced in comparison to age-matched controls (Fig. 2) by 6- to 30-fold in four independent experiments (Table 1). We therefore presume that rapid expansion of remaining CD4⁺CD8⁺ cells from anti-class I treated mice (see Fig. 1) in a MLC prevented the detection of a major difference in CTL precursor frequency in bulk cultures. This profound depletion of CTL precursors indicates that anti-class I treatment affected development of CD4⁺CD8⁺ cells by both phenotypical and functional criteria.

To determine whether anti-class I treated mice can generate any functional T cells, we analysed their ability to secrete IL-2 in response to MHC alloantigens. The levels of IL-2-producing precursor cells in spleen from control and treated mice were

Table 1 Frequency of cytotoxic T cells and IL-2-producing T cells in spleen from control and anti-class I treated mice

Experiment	C3H responder cells	Stimulator cells	Frequency of CTL precursors	Frequency of IL-2 producers
1*†	Control	BALB/c	1/4,000	1/250
	Treated	BALB/c	<1/10 ⁵	1/220
	Control	B10.A	1/15,000	
	Treated	B10.A	<1/2 × 10 ⁵	
2†	Control	BALB/c	1/15,000	1/1,400
	Treated	BALB/c	<1/2 × 10 ⁵	1/600
	Control	B10.A	1/20,000	
	Treated	B10.A	<1/2 × 10 ⁵	
3	Control	BALB/c	1/8,000	
	Treated	BALB/c	<1/5 × 10 ⁵	
4	Control	BALB/c	1/10,000	
	Treated	BALB/c	<1/60,000	

Responder cells were from C3H mice treated from birth with either saline or 15-1-5 mAb as indicated in Fig. 1. CTL-precursor frequency was determined as in Fig. 2 and that of IL-2-producing cells by a modification of a method described earlier¹³. Stimulator cells were nonirradiated resident peritoneal cells treated with anti-Thyl plus complement to eliminate T cells. Varying numbers of responder cells were cultured in 96-well microtitre plates with 5×10^4 stimulator cells in a final volume of 0.2 ml Dulbecco medium with 10% fetal calf serum (FCS), 2 mM L-glutamine, and 5×10^{-5} M 2-mercaptoethanol. After 5 d, supernatants were cultured with 3×10^3 CTLL cells and ³H-thymidine incorporation was measured after 24 h. Cultures were defined as positive when the radioactivity incorporated exceeded by >3 s.d. the mean ³H-thymidine incorporation of CTLL cells grown in control supernatants. In all experiments, a linear semilogarithmic relationship between the proportion of negative cultures and the dose of responder cells was observed. Minimal estimates of the precursor frequencies were calculated by plotting the % negative cultures against cell dose, according to the Poisson model, and by using the minimum χ^2 method of estimating the frequency.²¹

* Experiment also shown in Fig. 2.

† Experiments which also include class I MHC-region differences to illustrate that most of the CTL response is a reflection of a response to class I alloantigens.

comparable (Table 1). These experiments do not distinguish between class I and class II alloantigen-induced IL-2 production, but most of the IL-2 precursors detected in a MLC against a full MHC difference probably react against class II alloantigens¹³. These experiments point towards a selective effect of anti-class I treatment on the development of CTL, but not IL-2-producing precursor cells.

To control for possible side effects of the anti-class I mAb treatment, we first treated mice with similar amounts of an anti-class I mAb (34-1-2, anti-H-2^d, ref. 14) to an irrelevant class I type. No effects on cytotoxic T-cell function or T-cell subsets were observed (data not shown). Second, as most somatic cells, including T lymphocytes, express class I MHC molecules, it was possible that anti-class I coated lymphocytes were removed from the circulation by Fc receptor-bearing cells. It is difficult to explain how such a depletion might specifically affect CD4⁺CD8⁺ T cells but to rule out this possibility, we treated newborn (C3H × C57BL/6) F₁ mice with the anti-class I mAb (15-1-5). In such mice, where the injected mAb can react with C3H but not with C57BL/6 MHC antigens no effect on T-cell development was observed. Finally, adult, 6-week-old C3H mice were treated for 8 d with 15-1-5 mAb and, again, the number of CD4⁺CD8⁺ T cells was not reduced (Fig. 3). Taken together, these experiments strongly argue against the possibility that the anti-class I treatment resulted in selective removal of CD4⁺CD8⁺ cells by virtue of their higher sensitivity to anti-class I opsonization or by their more efficient removal. Possible inhibitory effects of anti-class I mAb carried over into the culture are also ruled out by two observations. First, treatment did not affect CTL responses in bulk MLC cultures, where the number of

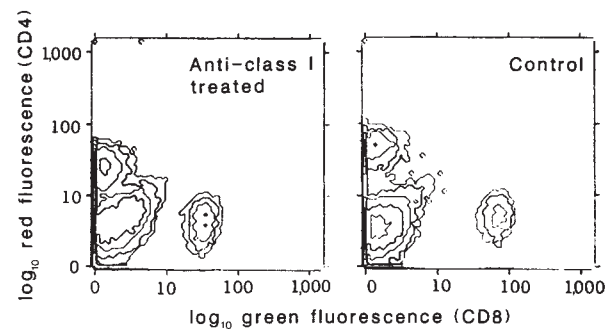


Fig. 3 Two-colour FCM analysis of CD4 and CD8 expression on spleen cells from adult control (right) and adult anti-class I treated (left) C3H mice.

Methods. 6-week-old C3H (H-2^k) mice were treated for 8 d with 500 µg of anti-H2K^d mAb 15-1-5 per g body weight as in Fig. 1. This dose saturates class I molecules in thymus and spleen (data not shown). Staining and FCM analysis was as in Fig. 1. Data are from three experiments. Per cent CD4⁺CD8⁺ cells: for control spleen, 8.81; for anti-class I spleen, 7.72. Per cent CD4⁺CD8⁺ cells: for control spleen, 9.97; for anti-class I spleen, 12.42.

responder cells is much higher (see above). Second, C3H spleen cells preincubated with 15-1-5 mAb under saturating conditions responded as well as control cells in the CTL-precursor frequency assay (data not shown). All of the data point, instead, to an indirect effect of the anti-class I treatment leading to the defect in development of functional CD4⁺CD8⁺ T cells.

It is not clear why the decrease in CTL-precursor frequency (Table 1) is generally more pronounced than the decrease in the number of CD4⁺CD8⁺ cells (Fig. 1). Some of the remaining CD4⁺CD8⁺ cells may belong to a nonfunctional subset, in keeping with the observation that not all thymic CD4⁺CD8⁺ cells (in contrast to CD4⁺CD8⁺ cells) express a T-cell receptor¹⁵. Whether the remaining CD4⁺CD8⁺ cells in anti-class I treated mice represent a subset developing independent of class I expression, or a subpopulation that has escaped the consequences of anti-class I treatment, is also not known. We favour the latter explanation, as it was difficult to achieve more than 95% saturation of the class I molecules with mAb.

We and others have previously demonstrated that class-I-restricted T cells express restriction specificities for both thymic and extrathymic MHC determinants, while class-II-specific T cells appear to undergo a strictly thymic-dependent differentiation process¹⁶⁻¹⁸. The results presented here indicate that the thymus is at least one of the sites where class-I MHC expression is essential for CD4⁺CD8⁺ T-cell development to occur.

Interestingly, although development of CD4⁺CD8⁺ and CD4⁺CD8⁺ cells can be inhibited by antibody to the appropriate class of MHC, the number of CD4⁺CD8⁺ T cells is normal in both anti-class I and anti-class II treated mice, suggesting that MHC interactions are not necessary for the generation of CD4⁺CD8⁺ cells. Administration of mAbs (KJ16) against TCR V_β8 prevents the appearance of mature T cells expressing receptors that contain V_β8, suggesting that TCR-ligand interactions are required for T-cell selection to occur¹⁹. A small subpopulation of CD4⁺CD8⁺, class-II-specific cytotoxic T cells develops normally in anti-class II treated animals²⁰, raising questions about whether interactions with MHC are necessary for development of cytotoxic T cells. The current data, however, unequivocally demonstrate a requirement for interactions with class I MHC for development of CD4⁺CD8⁺ T cells.

The simplest explanation for the present findings is that TCR-MHC interactions are also essential for the positive selection of class-I-restricted cytotoxic T cells. It is also possible, as recently suggested²⁰, that anti-class I and anti-class II mAbs block interaction of CD8 and CD4 molecules with class I and class II MHC determinants respectively, preventing expansion of CD4⁺CD8⁺ and CD4⁺CD8⁺ T cells. In either case, we con-

clude that interactions with self-MHC molecules are not required for either generation of CD4⁺CD8⁺ T cells or expression of the CD4 and CD8 antigens, but are essential for the development of both CD4⁺CD8⁺ and CD4⁺CD8⁻ subsets of mature T cells.

We thank Drs Alfred Singer, Jonathan D. Ashwell, Lucio Tentori and Ronald N. Germain for discussions, Ms Margaret A. Weston for technical support, Mrs Susan O. Sharrow for support and advice with the SCM experiments and Ms Ellie Irby for preparation of the manuscript. S. M-G. is an NIH Visiting Fellow from the Institute Ruder Boskovic, Zagreb, Croatia, Yugoslavia.

Received 24 February; accepted, 7 April 1988.

- Swain, S. L. *Immun. Rev.* **74**, 129-142 (1983).
- Dialynas, D. P. *et al. Immun. Rev.* **74**, 29-56 (1983).
- Zinkernagel, R. M., Callahan, G. N., Klein, J. & Dennert, G. *Nature* **271**, 251-253 (1978).
- Zinkernagel, R. M. *et al. J. exp. Med.* **147**, 882-896 (1978).
- Fink, P. J. & Bevan, M. J. *J. exp. Med.* **148**, 766-775 (1978).
- Kruisbeek, A. M., Hodes, R. J. & Singer, A. *J. exp. Med.* **153**, 13-29 (1981).
- Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
- Kruisbeek, A. M., Fultz, M. J., Sharrow, S. O., Singer, A. & Mond, J. J. *J. exp. Med.* **157**, 1932-1946 (1983).
- Kruisbeek, A. M. *et al. J. exp. Med.* **161**, 1029-1047 (1985).
- Kruisbeek, A. M., Sharrow, S. O., Mathieson, B. J. & Singer, A. *J. Immun.* **127**: 2168-2176, 1981.
- Ozato, K., Mayer, N. & Sachs, D. H. *J. Immun.* **124**, 533-540 (1980).
- Fowlkes, B. J., Edison, J., Mathieson, B. J. & Chused, T. M. *J. exp. Med.* **162**, 802-822 (1985).
- Ryser, J. E. & MacDonald, R. H. *J. Immun.* **122**, 1691-1696 (1979).
- Ozato, K., Mayer, N. M. & Sachs, D. H. *Transplantation* **34**, 113-120 (1982).
- Bluestone, J. A., Pardoll, D., Sharrow, S. O. & Fowlkes, B. J. *Nature* **326**, 82-84 (1987).
- Bradley, S. M., Kruisbeek, A. M. & Singer, A. *J. exp. Med.* **156**, 1650-1664 (1982).
- Kruisbeek, A. M., Davis, M. L., Matis, L. A. & Longo, D. L. *J. exp. Med.* **160**, 839-857 (1984).
- Kast, W. M., DeWaal, L. P. & Melief, C. J. M. *J. exp. Med.* **160**, 1752-1766 (1984).
- McDuffie, M., Born, P., Marrack & Kappler, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8728-8732 (1986).
- Mizuochi, T., Tentori, L., Sharrow, S. O., Kruisbeek, A. M. & Singer, A. *J. exp. Med.*, (in the press).
- Taswell, C. J. *Immun.* **126**, 1614-1619 (1981).

Insulin-regulatable tissues express a unique insulin-sensitive glucose transport protein

David E. James^{*†}, Robert Brown^{*}, Javier Navarro[†] & Paul F. Pilch^{*}

Departments of ^{*}Biochemistry and [†]Physiology, Boston University School of Medicine, 80 E. Concord Street, Boston, Massachusetts 02118, USA

At least three different glucose transport systems exist in mammalian cells. These are: (1) the constitutively active, facilitative carrier characteristic of human erythrocytes¹, Hep G2 (ref. 2) cells and rat brain³; (2) the Na-dependent active transporter of kidney and small intestine^{4,5}; and (3) the facilitative carrier of rat liver (B. Thorens and H. F. Lodish, personal communication). A fourth possible glucose transport system is the insulin-dependent carrier that may be specific to muscle and adipose tissue⁶. This transporter resides primarily in an intracellular compartment in resting cells from where it translocates to the cell surface upon cellular insulin exposure⁷⁻⁹. This raises the question of whether hormonal regulation of glucose transport is conferred by virtue of a tissue-specific signalling mechanism or a tissue-specific glucose transporter. Here we present data supporting the latter concept based upon a monoclonal antibody against the fat cell glucose transporter that identifies a unique, insulin-regulatable glucose transport protein in muscle and adipose tissue.

A major mechanism by which insulin stimulates glucose transport in rat adipocytes is by increasing the number of cell-surface transport proteins concomitant with the depletion of

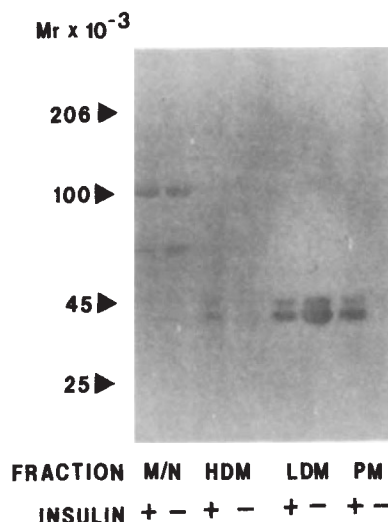


Fig. 1 Western blot analysis of adipocyte subcellular fractions using mAb 1F8. SDS-PAGE of subcellular membrane fractions (50 µg) obtained from adipocytes following incubation in the absence or presence of insulin (10^{-8} M) for 15 min: mitochondria/nuclei (M/N), high-density microsomes (HDM), low-density microsomes (LDM) and plasma membranes (PM). Mab 1F8, immunoblotted a doublet of 43 and 47K in the LDM of basal cells. Upon exposure of cells to insulin, the immunolabelled doublet shifted from LDM to PM. The higher-relative-molecular-mass bands labelled in the M/N fraction were also present in immunoblots using pre-immune mouse sera.

Methods. Rat adipocytes were isolated, and subcellular fractions were obtained using a differential centrifugation procedure as previously described¹⁸. Membranes were subjected to SDS-PAGE on a 5-13% resolving gel in the presence of 10 mM dithiothreitol using the system of Laemmli²⁷. Proteins were transferred to nitrocellulose which was blocked, and incubated with Protein-A affinity-purified mAb 1F8 at a dilution of 1/250. Immune complexes were visualized with biotinylated sheep anti-mouse antisera followed by incubation with horse radish peroxidase linked to avidin (Amersham, diluted 1/100), using 4-chloronaphthol as substrate. Mab 1F8 was produced following immunization of Balb-C mice with partially purified intracellular glucose transporter-containing vesicles (50 µg intra-peritoneally in MPL + TDM emulsion, Rib Immuno Chem.) obtained from rat adipocytes as previously described¹². Mice were boosted twice with purified vesicles, as above, at 3-week intervals and once (25 µg of purified vesicles intravenously) 4 days before fusion of spleen cells with SP2 mouse myeloma cells, as previously described²⁸. Dot blot analysis, using 1 µg of sonicated sucrose-gradient purified transporter-containing vesicles per well, was used as the initial screen. Positive clones were subsequently screened via a differential western blot, using preparative SDS-PAGE. This was done using plasma membranes isolated from cells incubated in the absence or presence of insulin (10^{-8} M). Clones which exhibited a differential response were subsequently subcloned by limiting dilution. Nine different subclones of 1F8 were obtained, all of which exhibited an identical immunolabelling pattern to that shown above.

transporters from an intracellular pool, called low-density microsomes⁷⁻¹¹. We have recently described a procedure for partially purifying the intracellular vesicles that comprise the translocatable pool of glucose transporters¹². Here, we have used these partially purified vesicles as an immunogen to produce antibodies to proteins whose cellular location is dynamically regulated by insulin. We obtained a monoclonal antibody (mAb), 1F8, that immunolabels a doublet of the size reported for the adipocyte glucose transporter^{13,14}, in the low-density microsomes from resting cells (Fig. 1). Consistent with the expected translocation mechanism, there is no detectable immunoblotting in the plasma membrane fraction from resting cells, but upon insulin exposure, there is a decrease ($\approx 50\%$) in the immunolabelled doublet from the intracellular pool concomitant with a corresponding increase in the cell surface label

[†] Present address: Department of Cell Biology and Physiology, Washington University School of Medicine, Box 8101, 660 South Euclid Avenue, St. Louis, Missouri 63110, USA.

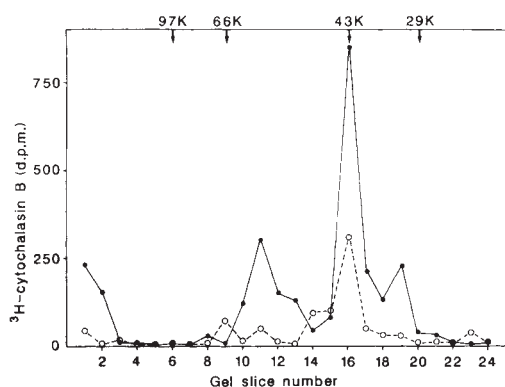


Fig. 2 Immunoabsorption of solubilized adipocyte microsomes, photoaffinity-labelled with [^3H]cytochalasin B, using mAb 1F8. LDMs were resuspended in 10 mM sodium phosphate buffer, pH 7.4. Membranes (1 mg) were photoaffinity-labelled with [^3H]cytochalasin B (1 μM) in the absence and presence of 0.6 M D-glucose, as described previously²⁹. Membranes were immediately pelleted at 164,000 g for 90 min at 4°C, sonicated in lysis buffer containing 1% Triton X-100, 0.5% NP-40, 1 mM EDTA and 1% SDS in phosphate-buffered saline (PBS), pH 7.4, for 30 s and incubated for 30 min at 37°C. Solubilized membranes were diluted (1/10) with lysis buffer, prepared as described above, excluding SDS. Ascites fluid (5 μl) containing mAb 1F8 was incubated with goat anti-mouse antisera linked to Sepharose beads (50 μl , Cappel) for 1 h at 37°C. Beads were washed three times with PBS and incubated overnight, with shaking at 4°C with solubilized membrane proteins. The beads were pelleted and washed four times with PBS containing 2% Triton X-100, 1% NP-40 and 2 mM EDTA, pH 7.4, and a further two times with PBS. The washed beads were boiled for 2 min in the presence of Laemmli sample buffer containing 10 mM dithiothreitol and subjected to SDS-PAGE. Gel lanes were cut into 0.3-cm slices, using a gel slicer, boiled in 30% hydrogen peroxide and counted. Shown is the radioactivity in each gel fraction (measured as disintegrations per min (d.p.m.), for membranes photoaffinity-labelled with [^3H]cytochalasin B in the absence (●) or presence of glucose (○). Immunoabsorptions were also performed, as described above, using pre-immune mouse serum (100 μl). Quantitation of the SDS-PAGE fractions from these samples revealed no significant difference from the background of the scintillation counter. Pre-stained relative-molecular-mass markers (Bethesda) were electrophoresed in an adjoining lane of the gel (their relative migration profile is denoted at the top of the figure).

(Fig. 1). The doublet is a variable result (Fig. 4), also seen by others studying glucose transporters¹⁵, that appears to be an artefact of sample preparation for sodium-dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).

We obtained definitive evidence that 1F8 is specific for the adipocyte glucose transporter by immunoprecipitating adipocyte microsomes that were photoaffinity-labelled with [^3H]cytochalasin B, a specific marker of glucose transporters in a variety of tissues¹⁶⁻¹⁸. Figure 2 shows that 1F8 immunoprecipitates a protein with a relative molecular mass (M_r) of 43,000 (43K) that is labelled with cytochalasin B in a D-glucose-displaceable way. Pre-immune mouse serum did not precipitate label (data not shown). The smaller peak of label seen at ~57K has been seen by others using cytochalasin photolabelling protocols¹⁷, and may be due to aberrant behaviour of transporters in SDS-PAGE.

Previous studies that have used antibodies to determine the relative concentration of insulin-regulatable glucose transporters from different membrane fractions of adipocytes have shown that insulin increases the number of plasma membrane glucose transporters 2-3-fold^{13,14,19-21}. In contrast, insulin causes a 5-10-fold increase in plasma membrane cytochalasin-B binding sites^{7,21}. To resolve this discrepancy, we examined the relationship between transporter number assayed by cytochalasin-B binding and by immunoblotting with 1F8 in subcellular fractions

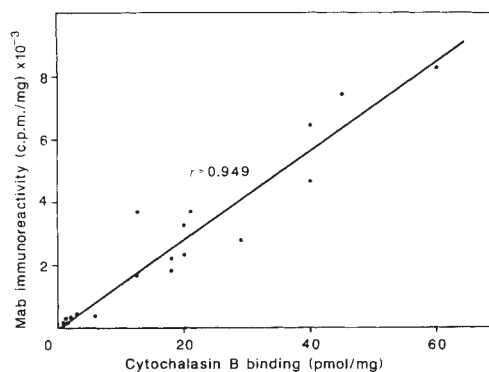


Fig. 3 Correlation of glucose transporter concentration in adipocyte subcellular fractions measured by either cytochalasin-B binding or immunoblotting with mAb 1F8. LDM and PM were obtained from adipocytes, incubated in the presence or absence of insulin (10^{-8} M), as described in Fig. 1. Cytochalasin-B binding isotherms were compiled for each membrane sample⁷ and the B_{max} for the binding was estimated. Each membrane sample was also subjected to SDS-PAGE, in duplicate, followed by immunoblotting using mAb 1F8. Immune complexes were visualized by autoradiography following incubation of nitrocellulose with [^{125}I]labelled goat anti-mouse antisera (20,000 counts per minute (c.p.m.) ml^{-1}), as previously described. The immunolabelled bands, of 43K, were excised from the nitrocellulose and counted. Correction was made for background by counting non-labelled areas of nitrocellulose. The correlation was performed using linear regression analysis. The correlation coefficient (r) and the computer fitted line are shown. The detection of transporter in the PM fraction of unstimulated cells by antibody is at the limit of detection, and thus our fold-stimulation by insulin represents a minimal estimate.

from resting and insulin-treated cells (Fig. 3). The increase in plasma membrane transporters due to insulin was 6-10-fold in the three experiments performed, and the two procedures used to measure transporters gave data that were highly correlated ($r = 0.949$) supporting our original contention concerning the specificity of 1F8.

Antibody 1F8 immunolabels glucose transporters only from insulin-sensitive tissues, namely muscle and fat (Fig. 4, bottom panel). Transporter was not detected by 1F8 in rat brain or human erythrocytes. In contrast, an anti-human erythrocyte glucose transporter polyclonal sera immunolabelled transporter from these two sources (Fig. 4, top panel). Neither antibody was immunoreactive with rat erythrocytes, kidney, jejunum and liver.

These data (Fig. 4) show that insulin sensitive tissues express a glucose transporter that is antigenically unique as compared to previously described transport proteins. Others have also suggested that adipocytes might contain a unique transporter based on several lines of evidence. The kinetics of glucose flux in adipocytes are symmetric, but are asymmetric in erythrocytes²². There are also substrate-specificity differences between these two systems²³. The fat cell low-density microsomes have been reported to contain two forms of transporter as determined by photolabelling with cytochalasin B and isoelectric focusing²⁴. Only one of these forms responded to insulin. A recent study using anti-human erythrocyte transporter antibody to characterize the immunoreactivity of fat cell transporters also reached the conclusion that fat cells have at least two transporter types differing in their immunoreactivity with the antiserum used¹⁹.

Taken together, our data and the data of others²²⁻²⁴ strongly support the notion that cellular insulin sensitivity is conferred by the tissue-specific expression of a unique glucose transporter. It is probable that insulin-sensitive tissues express a 'constitutive' transporter as well, because many polyclonal sera recognize a transport protein in the plasma membrane of resting cells^{13,14,18-21}, whereas 1F8 does not (Fig. 1). It is not yet clear

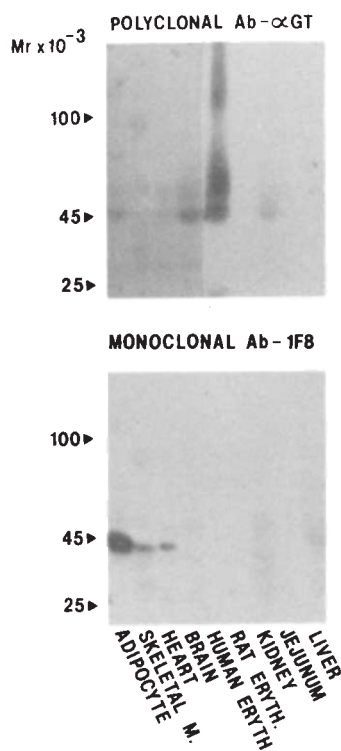


Fig. 4 Tissue specificity of mAb 1F8 and an anti-human erythrocyte glucose transporter antisera. Microsomes were isolated from rat heart, brain, liver, jejunum, kidney and skeletal muscle by homogenizing tissues in a buffer containing 20 mM HEPES, 1 mM EDTA and 250 mM sucrose, pH 7.4 for 10–20 s using a polytron. Homogenates were centrifuged at 15,000g for 20 min at 4 °C. The supernatants from heart and skeletal muscle homogenates were incubated in the presence of 0.8 M KCl for 30 min at 4 °C. Microsomes from each tissue were obtained by centrifugation of the low speed supernatants at 160,000g for 90 min at 4 °C. In the case of adipocytes, low-density microsomes were used. For rat and human erythrocytes, ghosts were prepared. Membranes (200 µg) were solubilized in Laemmli sample buffer and subjected to SDS-PAGE. Only 5 µg of human erythrocyte ghost protein was used. Immunoblotting was performed, as described in Fig. 1, except that iodo goat-anti mouse and iodo goat anti-rabbit IgG antisera (Dupont, NEN) were used to identify specific immunostaining with mAb 1F8 and anti-human erythrocyte glucose transporter antisera, respectively.

whether the various polyclonal sera recognize the insulin regulatable transporter, albeit with lower affinity, or whether the two species are unrelated. Three possibilities exist as to the relationship of the two postulated adipocyte glucose transporters. First, they may be the product of separate genes, both expressed in the same cell, as in renal and intestinal cells where two transporter types are apparently made^{2–5,25}. Second, they may be the product of the same gene where alternative messenger RNA splicing has occurred²⁶, or third, they may differ due to post-translational modification of the same gene product. It is known only that the epitope recognized by 1F8 is facing the cell cytoplasm because immobilized antibody can immunoadsorb transporter-containing vesicles (data not shown). Thus, it is necessary to characterize the insulin-sensitive transporter at the genetic level and to purify and characterize amounts large enough to describe the biochemical feature(s) that account for its unique antigenicity and tissue distribution.

The present data suggest that tissue-specific, insulin-regulated glucose transport may be conferred by virtue of the tissue-specific expression of an insulin-regulatable transport protein. This concept introduces a potential mechanism whereby specific abnormalities in insulin-dependent glucose utilization may

occur in certain disease states, in particular, type-II diabetes. Further studies are necessary to define biological situations where possible modifications in the insulin-regulatable glucose transporter, or its environment, may compromise insulin action and/or contribute to the phenotype of type-II diabetes.

This work was supported by the NIH and the NSF. D.E.J. is the recipient of a Juvenile Diabetes Foundation International, postdoctoral fellowship. We thank Lisa DeAngelo for technical assistance, Kerri James for artwork and Dr Chung Wang, Harvard University, Cambridge, MA, for his gift of anti-human erythrocyte glucose transporter antibody.

Received 8 December 1987; accepted 29 March 1988.

1. Jung, C. in *The Red Blood Cell*, Vol. II (ed. Surgenor, D. M.) 705–751 (Academic, New York, 1975).
2. Mueckler, M. *et al. Science* **229**, 941–945 (1985).
3. Birnbaum, M. J., Haspel, H. C. & Rosen, O. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5784–5788 (1986).
4. Wu, J.-S. R. & Lever, J. E. *Biochemistry* **26**, 5958–5962 (1987).
5. Hediger, M. A., Coady, M. J., Ikeda, T. S. & Wright, E. M. *Nature* **330**, 379–381 (1987).
6. Elbrink, J. & Bihler, I. *Science* **188**, 1177–1184 (1975).
7. Cushman, S. W. & Wardzala, L. J. *J. biol. Chem.* **255**, 4758–4762 (1980).
8. Suzuki, K. & Kono, T. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2542–2545 (1980).
9. Blok, J., Gibbs, E. M., Lienhard, G. E., Slot, J. W. & Geuze, H. J. *J. cell Biol.* **106**, 69–76 (1988).
10. Lienhard, G. E., Kim, H. H., Ransome, K. J. & Gorga, J. C. *Biochem. biophys. Res. Commun.* **105**, 1150–1156 (1982).
11. Oka, Y. & Czech, M. P. *J. biol. Chem.* **259**, 8125–8133 (1984).
12. James, D. E., Lederman, L. & Pilch, P. F. *J. biol. Chem.* **262**, 11817–11824 (1987).
13. Shanahan, M. F., Olson, S. A., Weber, M. J., Lienhard, G. E. & Gorga, J. C. *Biochem. biophys. Res. Commun.* **107**, 38–43 (1982).
14. Wheeler, T. J., Simpson, I. A., Sogin, D. C., Hinkle, P. C. & Cushman, S. W. *Biochem. biophys. Res. Commun.* **105**, 89–95 (1982).
15. Haspel, H. C., Birnbaum, M. J., Wilk, E. W. & Rosen, O. M. *J. biol. Chem.* **260**, 7219–7225 (1985).
16. Shanahan, M. F. *J. biol. Chem.* **257**, 7290–7293 (1982).
17. Pessin, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2286–2290 (1982).
18. Simpson, I. A. & Cushman, S. W. *A. Rev. Biochem.* **55**, 1059–1089 (1986).
19. Wang, C. *J. biol. Chem.* **262**, 15689–15695 (1987).
20. Haspel, H. C., Rosenfeld, M. G. & Rosen, O. M. *J. biol. Chem.* **263**, 398–404 (1988).
21. Joost, H. G., Weber, T. M. & Cushman, S. W. *Biochem. J.* **249**, 155–161 (1988).
22. Taylor, L. P. & Holman, G. D. *Biochim. biophys. Acta* **642**, 325–335 (1981).
23. Holman, G. D. & Rees, W. D. *Biochim. biophys. Acta* **685**, 78–86 (1982).
24. Horuk, R., Matthei, S., Olefsky, J. M., Baly, D., Cushman, S. W. & Simpson, I. A. *J. biol. Chem.* **261**, 1823–1828 (1986).
25. Kimmich, G. A. & Randle, J. *Am. J. Physiol.* **241**, C227–C232 (1981).
26. Breitbart, R. E., Andreadis, A. & Nadal-Ginard, B. *A. Rev. Biochem.* **56**, 467–495 (1987).
27. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
28. Galfre, G. & Milstein, C. *Meth. Enzym.* **73**, 3–46 (1981).
29. Wessling, M. & Pilch, P. F. *Biochim. biophys. Acta* **777**, 123–132 (1984).

The N-terminal region of the chicken progesterone receptor specifies target gene activation

Laszlo Tora, Hinrich Gronemeyer, Bernard Turcotte, Marie-Pierre Gaub & Pierre Chambon*

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 11 rue Humann 67085 Strasbourg Cedex, France

Steroid hormone receptors belong to a family of nuclear receptors that trigger transcriptional activation of target genes by specific binding to DNA recognition sequences, usually located in the 5'-flanking region of the target gene. Nuclear receptors appear to be segmented proteins¹ and extensive structure-function analyses have attempted to elucidate the functional significance of individual segments. Two of these regions have been defined as the domains responsible for recognition of responsive elements of target genes (region C) and hormone binding (region E) (refs 2–7). But the functional significance of the N-terminal region (A/B), which diverges extensively even for a given receptor between different species, has remained obscure. We have previously cloned,

* To whom correspondence should be addressed.

expressed and analysed the chicken progesterone receptor (cPR) (ref. 8). This receptor and its human homologue from T47D breast cancer cells are unique among the steroid hormone receptors in that two forms, A and B, are present in equal amounts in cytosolic extracts, the latter having the higher molecular weight^{9,10}. For the chicken progesterone receptor, we have presented evidence suggesting that the cPR form A corresponds to an N-terminally truncated form of B (ref. 8). Here we report on the functional difference between the forms A and B in the transcriptional activation of two target genes.

By expressing the 786-amino acid chicken progesterone receptor from the complementary DNA expression vector cPR1, we have shown that it is indistinguishable from the cPR form B observed in cytosolic extracts of chick oviduct⁸. There is only one methionine present in the N-terminal half of cPR form B (Met 128, see Fig. 1a and ref. 8). Initiating translation from this methionine using the vector cPR2 yields a stable protein indistinguishable from the oviduct cPR form A in size, immunoreactivity and hormone binding⁸. Thus, it appears that the cPR form A corresponds to form B without the N-terminal 127 amino acids (Fig. 1a and ref. 8). When HeLa cells are co-transfected with the expression vectors cPR1 or cPR2 and a reporter gene containing the hormone responsive element (HRE) present in the mouse mammary tumour virus (MMTV) long terminal repeat (LTR), both cPR forms transactivate transcription efficiently, with form B being about twice as efficient as form A (ref. 8). In contrast, no activation of transcription was detected in HeLa cells using either cPR1 or cPR2 with reporter genes in which fragments of the chicken ovalbumin promoter region from -1,348 to +1 were positioned upstream of the promoterless bacterial chloramphenicol transferase (CAT) gene (data not shown, see also below). Similar results were obtained with cPR1 and ovalbumin-CAT reporter genes when primary chicken embryo fibroblast (CHEF) cultures were transiently transfected. But by expressing the cPR form A from the corresponding vector cPR2, progestin-dependent stimulation of CAT activity was readily detected (see below). Both the cPR forms A and B efficiently activated transcription in CHEF cells from co-transfected MMTV-CAT (see below). By progressively deleting ovalbumin promoter 5'-sequences, a progestin response element (PRE) could be located in a 59-base pair region (Fig. 1b) between positions -58 and +1 of the chicken ovalbumin gene promoter¹¹ (L.T., unpublished results and see below). The failure of cPR1 to stimulate transcription of the ovalbumin-CAT reporter gene was not due to a degradation of form B in CHEF cells, because both cPR1 and cPR2 generated similar amounts of stable receptor proteins capable of binding tightly to CHEF cell nuclei, as analysed by hormone binding studies and Western blotting (not shown). From these results we concluded that (1) the two forms of the cPR can activate the HRE present in the MMTV-LTR, both in HeLa or CHEF cells, (2) activation of the ovalbumin gene PRE requires a cell-specific factor absent in HeLa cells in addition to the progesterone receptor and (3) this PRE is apparently only activated by the N-terminally truncated form A of the cPR.

To support these conclusions we analysed both MMTV and ovalbumin reporter genes simultaneously in the same transfection for transcriptional activation by the cPR forms A or B. We used quantitative S1 nuclease analysis in parallel to assay ovalbumin gene transcription, and CAT-assays to monitor activation of the MMTV-HRE. We also included a double internal control system by co-transfecting pG1B (ref. 8) to normalize the signals of S1 nuclease analysis and the β -galactosidase expression vector pCH110 to normalize extracts for CAT-assay⁸. For S1 nuclease analysis, an OV-GLOB reporter gene was constructed containing the ovalbumin gene promoter sequences from -58 to +1 in front of the promoterless rabbit β -globin gene, as illustrated in Fig. 1c. When analysed by S1 nuclease mapping, RNA transcripts initiating at +1 of OV-GLOB yield a protected fragment of 152 nucleotides (illustrated in Fig. 1c), whereas

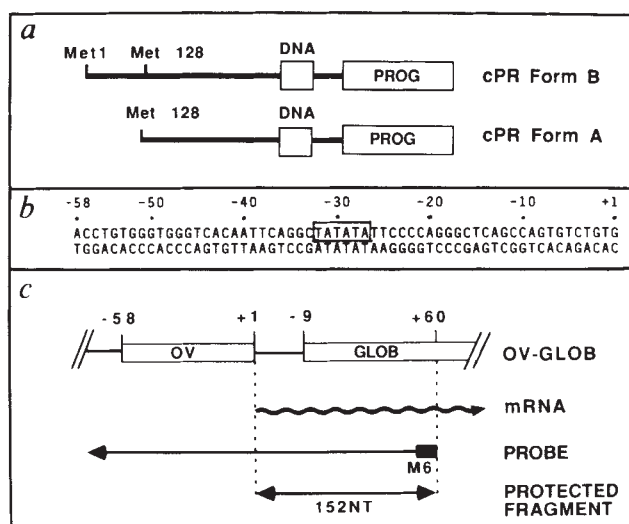


Fig. 1 a, Alignment of forms A and B of the chicken progesterone receptor (cPR). The form with lower relative molecular mass (cPR form A) corresponds to an N-terminally truncated version of the full-length receptor form B (ref. 8). The putative N-terminus of the natural form A, Met-128, is indicated, as well as the DNA and hormone (PROG) binding domains. b, Sequence of the chicken ovalbumin gene promoter region used for the construction of the OV-GLOB reporter gene. The G at position +1 replaces A in the wild-type sequence¹¹. c, Schematic illustration of S1 nuclease mapping analysis with OV-GLOB. The DNA between +1 of OV and -9 of GLOB corresponds to the SV40 sequence from position 5,171 to 5,227 (ref. 27) and a piece of polylinker resulting from the cloning strategy used. This sequence is inert in the functional assays described in this report. M6 is a synthetic oligonucleotide¹² complementary to position +39 to +60 of the rabbit β -globin gene used to generate the Probe by primer extension⁸. After hybridization to RNA initiating at +1 of OV-GLOB and treatment with S1 nuclease, the Probe yields a protected fragment of 152 nucleotides.

initiation at +1 of β -globin in the internal control recombinant pG1B results in a fragment of 60 nucleotides⁸. Both OV-GLOB and MMTV-CAT were co-transfected with the respective control recombinants and cPR1 or cPR2 in the presence or absence of 10^{-7} M progesterone. After 48 hours, the cells were split into two fractions which were subjected to CAT-activity assay and S1 nuclease analysis. The results of the CAT-assay showed that, in CHEF cells, both forms of the chicken progesterone receptor activate the MMTV-HRE in the presence of progesterone, although form A is ≈ 5 times less efficient than form B (Fig. 2 top, lanes 1-4). Identical results, namely 19-fold stimulation by cPR form B and 4-fold stimulation by form A, were obtained when the OV-GLOB and pG1B recombinants were omitted from the transfection (not shown). In contrast, in the same cells where cPR1 efficiently activated the MMTV-HRE, no significant transactivation of transcription by the cPR form B was detected from OV-GLOB (Fig. 2 bottom, lanes 1 and 2; see also Fig. 3, lanes 4 and 5). When expressing the cPR form A, however, a clear stimulation of transcription was apparent in the presence of progesterone (Fig. 2 bottom, lanes 3 and 4, and Fig. 3, lanes 7 and 8; an average stimulation of 4.4-fold was determined from three different experiments).

The promoter region of the chicken ovalbumin gene that has been used to generate OV-GLOB (Fig. 1b) also contains an oestrogen-responsive element (L.T., M.P.G., A. Dierich, M. Bellard and P.C., in preparation), and exposure to oestradiol of CHEF cells transiently co-transfected with OV-GLOB and the human oestrogen receptor expression vector HEO (refs 4, 13) resulted in a 4- to 5-fold stimulation of transcription (Fig. 3, compare lanes 2 and 3). The close vicinity of the DNA

elements in the ovalbumin reporter gene that are recognized by the oestrogen and progesterone receptors prompted us to study whether the two activation systems function independently. Interestingly, we observed a two- to three-fold decrease of the oestrogen-induced transcription from OV-GLOB when cPR1 was co-transfected (Fig. 3, compare lanes 3 and 6), suggesting that the inability of the cPR form B to induce ovalbumin gene transcription was not just a result of an inability to recognize the ovalbumin gene promoter sequences. Moreover, no effect was observed in the absence of progesterone (not shown). The opposite was observed when cPR2 was transfected together with HEO and OV-GLOB. Densitometric determination of the signals as shown in lanes 3, 8 and 9, from three independent S1 nuclease mapping analyses reveals 4.2, 4.4 and 7.8-fold stimulations for OV-GLOB transcription by the oestrogen receptor alone, cPR form A alone and both receptors together, respectively, in the presence of the relevant hormones. Thus, the cPR form A and the oestrogen receptor have an additive stimulatory effect on ovalbumin gene transcription.

Several novel aspects of steroid hormone controlled gene transcription are shown by these experiments. Firstly, the observation that the cPR is able to stimulate ovalbumin gene transcription in CHEF cells, but not in HeLa cells, indicates that cell-specific factor(s) are indispensable for this activation. As only one form of the cPR can activate the ovalbumin gene PRE and as both forms in the same cells activate transcription from the MMTV-PRE, it is likely that these transcription factors differentially cooperate with the receptor forms. It will be interesting to define whether similar cell-specific factor requirements exist for other cloned progesterone receptor target genes from chicken¹⁴⁻¹⁶ or other species¹⁷.

Secondly, the N-terminal domain of the cPR apparently plays a specific role in the transcriptional activation. Previous work using glucocorticoid receptor mutants has demonstrated the requirement for an intact A/B domain to activate the MMTV-LTR-HRE efficiently^{18,19}. Until now, no further GREs have been analysed with respect to the functional role of the N-terminal domain. For the oestrogen receptor, however, a differential sensitivity of target genes towards deletions in the A/B domain has been observed: the *Xenopus* vitellogenin A2 gene oestrogen responsive element (ERE) is activated to a similar extent by the wild-type oestrogen receptor or by its N-terminally truncated mutant, whereas only 10-17% activation of the wild-type level is found for the pS2 gene ERE (ref. 3). Our present results clearly show that the N-terminal domain of the cPR is implicated in both efficiency and specificity of transcriptional transactivation. We report for the first time that the transcription of a responsive gene can be stimulated by the naturally occurring N-terminally truncated progesterone receptor form, and not by the intact form of the receptor. Interestingly, the human glucocorticoid receptor expression vector HG1 (ref. 3) can activate transcription in CHEF cells from OV-GLOB in the presence of glucocorticoids (not shown). The DNA-binding domain of the glucocorticoid receptor is very similar to that of the cPR (ref. 20), whereas the N-terminal regions are different. It is thus possible that the PRE of the ovalbumin gene promoter is recognized by both the progesterone and glucocorticoid receptors, as is the case for the MMTV-HRE. Thus it appears that it is the N-terminal 127-amino acid region of the cPR which is selectively responsible for the lack of activation of the ovalbumin gene PRE by the cPR form B. We have previously noted that this region contains a peculiar polyglutamic acid region⁸. A set of N-terminal chimeras and truncated receptors are currently being constructed to find out whether this polyglutamic acid stretch is the cause of the inability of cPR form B to activate transcription from the ovalbumin promoter.

At present it is unclear whether the differential activation of HREs by two products derived from the same receptor gene is a specific characteristic of the progesterone receptor. We are not aware of the existence of stable and functional forms similar

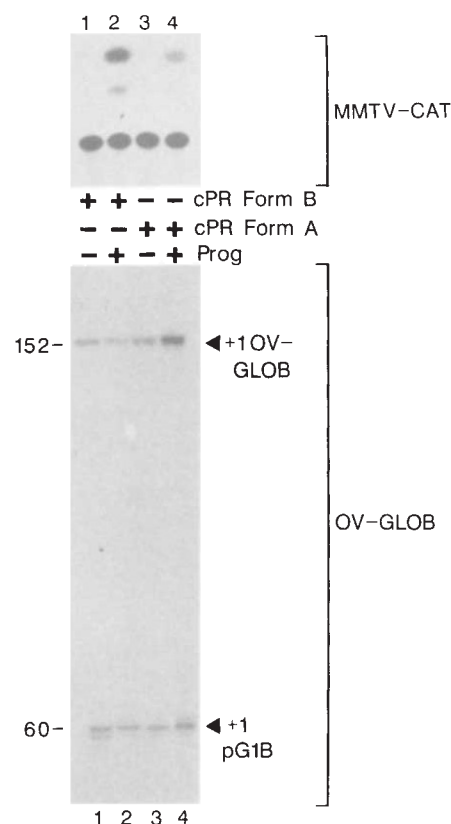


Fig. 2 CAT assay (top) and S1-nuclease mapping analysis (bottom) performed with cell extracts and cytosolic RNA, respectively, of the same transiently transfected CHEF cells, using the expression vectors cPR1 (form B) or cPR2 (form A) (see ref. 8) in the presence (+) or absence (-) of 10^{-7} M progesterone (Prog). Initiation of transcription at +1 in OV-GLOB (Fig. 1) results in the appearance of a fragment 152 nucleotides long (arrowhead +1 OV-GLOB) that is protected from S1 nuclease digestion. The internal control recombinant pG1B (ref. 8) yields a protected fragment of 60 nucleotides (+1 pG1B).

Methods. CHEF cells were prepared from 9-10-day-old chicken embryos²⁸ and grown in Dulbecco's modified Eagle's medium (DMEM) without phenol red and containing 5% charcoal-stripped fetal calf serum and 5% bacto-tryptone. After two passages, transfection was carried out with 10 μ g OV-GLOB, 1 μ g cPR1 or cPR2 (8), 0.4 μ g pG1B, 1 μ g MMTV-CAT (ref. 29), 3 μ g pCH110 and Bluescribe plasmid carrier to 20 μ g total DNA per $3-5 \times 10^6$ cells per 9 cm Petri dish, using the standard calcium phosphate coprecipitation technique³⁰. After 48 hours, cells were collected and resuspended in 300 μ l 10 mM Tris-HCl, pH 7.9, 10 mM NaCl, 2 mM $MgCl_2$ per plate, from which 200 μ l were used to prepare cytoplasmic RNA⁸, and the remainder served to generate cell extracts for CAT assays using standard protocols³¹ with β -galactosidase calibration to compensate for variations in transfection efficiencies⁸. Single-stranded DNA probe for S1 nuclease mapping was prepared by first hybridizing the 32 P-5'-labelled oligonucleotide M6 (Fig. 1b) to single-stranded M13mp18 carrying a 1.3-kb *Eco*RI fragment of OV-GLOB. After primer extension, the single-stranded probe was purified by polyacrylamide gel electrophoresis and electroelution. Hybridizations (68 $^{\circ}C$, 12 h) were carried out with 10-15 μ g RNA dissolved in 20 μ l 10 mM PIPES, pH 6.5, 400 mM NaCl and 50,000 c.p.m. single-stranded probe (excess probe). The samples were diluted to 200 μ l with 30 mM sodium acetate, pH 4.5, 3 mM $ZnCl_2$, 400 mM NaCl and digested with 100 units S1 nuclease (Appligene) for 1 h at 25 $^{\circ}C$. Nuclease-resistant hybrids were analysed on 8% acrylamide/8.3 M urea sequencing gels. The signals obtained on autoradiographs from three different independent experiments using different expression vector preparations were quantified by densitometric measurements (Hoefer GS300, GS350 Data System) and corrected for variations in transfection efficiency using the signals corresponding to pG1B transcription (+1 pG1B).

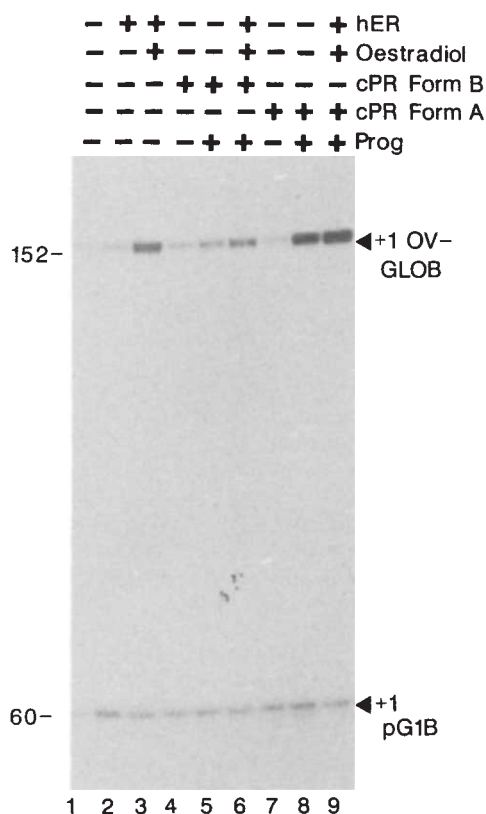


Fig. 3 cPR form B partially suppresses the human oestrogen receptor (hER)-mediated transcriptional induction from OV-GLOB, whereas cPR form A causes an additive stimulation, 10 μ g OV-GLOB, 0.4 μ g pG1B and the human oestrogen receptor expression vector HEO (refs 4, 13) were transiently co-transfected into CHEF cells together with cPR1 or cPR2 in the presence (+) or absence (-) of 10^{-8} M oestradiol and/or 10^{-7} M progesterone (Prog), as depicted at the top of the figure; transfection of 1 μ g hormone receptor expression vector DNA per $3-5 \times 10^6$ cells per 9 cm Petri dish is indicated by (+). Transfections of HEO, cPR1 and cPR2 alone, in the presence or absence of the corresponding hormones, as well as a mock transfection, are shown for comparison on the same S1-nuclease-mapping gel. Individual lanes are numbered at the bottom. Densitometric determination of the signals as shown in lanes 3, 6, 8 and 9 indicated stimulation of 3.8, 1.7, 6.4 and 7.8-fold, respectively, after correction for transcription from the internal control pG1B.

to the two cPR forms among other members of the steroid/thyroid/retinoic acid receptor gene family. In the case of the thyroid²¹⁻²³ and retinoic acid receptors²⁴⁻²⁶ we note, however, that at least two proteins respond to the same ligand, although they are derived from different genes, and have highly homologous DNA and steroid binding domains. Target gene specificity of these receptors could be modulated by the non-homologous N- and C-terminal regions in the same way as we have shown here for the two forms of the progesterone receptor.

We thank our colleagues of the LGME receptor group for a gift of the recombinant HEO as well as for challenging discussions, and S. Green and G. Richards for critically reviewing this manuscript. We are grateful to I. Scheuer and M. Digelmann for technical assistance, to C. Werle and B. Boulay for help in the construction and reproduction of the figures, and to the secretarial staff for typing this manuscript. B.T. is a postdoctoral fellow of the Medical Research Council of Canada. This work was supported by grants from the CNRS, the INSERM, the Fondation pour la Recherche Médicale, the Ministère de la Recherche et de l'Enseignement Supérieur and the Association pour la Recherche sur le Cancer and the Ligue Nationale contre le Cancer.

Received 21 March; accepted 15 April 1988.

- Green, S. & Chambon, P. *Nature* **324**, 18-25 (1986).
- Green, S. & Chambon, P. *Nature* **325**, 75-78 (1987).
- Kumar, V., Green, S., Stack, G., Berry, M., Jin, J. R. & Chambon, P. *Cell* **51**, 941-951 (1987).
- Kumar, V., Green, S., Staub, A. & Chambon, P. *EMBO J.* **5**, 2231-2236 (1986).
- Hollenberg, S. M., Giguère, V., Segui, P. & Evans, R. M. *Cell* **49**, 39-46 (1987).
- Godowski, P. J., Rusconi, S., Miesfeld, R. & Yamamoto, K. R. *Nature* **325**, 365-368 (1987).
- Danielson, M., Northrop, J. P. & Ringold, G. M. *EMBO J.* **5**, 2513-2522 (1986).
- Gronemeyer, H. *et al.* *EMBO J.* **6**, 3985-3994 (1987).
- Gronemeyer, H., Harry, P. & Chambon, P. *FEBS Lett.* **156**, 287-292 (1983).
- Wei, L. L. *et al.* *Biochemistry* **26**, 6262-6272 (1987).
- Gannon, F. *et al.* *Nature* **278**, 428-434 (1979).
- Zenke, M. *et al.* *EMBO J.* **5**, 387-397 (1986).
- Green, S. *et al.* *Nature* **320**, 134-139 (1986).
- Renkawitz, R., Schütz, G., von der Ahe, D. & Beato, M. *Cell* **37**, 503-510 (1984).
- von der Ahe, D., Renoir, J. M., Buchou, T., Baulieu, E. E. & Beato, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2817-2821 (1986).
- Gope, M. L. *et al.* *Nucleic Acids Res.* **15**, 3595-3607 (1987).
- Bailly, A., Le Page, C., Rauch, M. & Milgrom, E. *EMBO J.* **5**, 3235-3241 (1986).
- Hollenberg, S. M., Giguère, V., Segui, P. & Evans, R. M. *Cell* **49**, 39-46 (1987).
- Miesfeld, R., Godowski, P. J., Maler, B. A. & Yamamoto, K. R. *Science* **236**, 423-427 (1987).
- Jeltsch, J. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 5424-5428 (1986).
- Weinberger, C. *et al.* *Nature* **324**, 18-25 (1986).
- Thompson, C. C., Weinberger, C., Lebo, R. & Evans, R. M. *Science* **237**, 1610-1614 (1987).
- Benbrook, D. & Pfahl, M. *Science* **238**, 788-791 (1987).
- Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. *Nature* **330**, 444-450 (1987).
- Giguère, V., Ong, E. S., Segui, P. & Evans, R. M. *Nature* **330**, 624-629 (1987).
- Brand, N. *et al.* *Nature* (in the press).
- Tooze, J. *Molecular Biology of Tumor Viruses* 2nd Edn, Part 2 (Cold Spring Harbor Laboratory Press, 1982).
- Dierich, A., Gaub, M. P., Le Pennec, J. P., Astinotti, D. & Chambon, P. *EMBO J.* **6**, 2305-2312 (1987).
- Cato, A. C. B., Henderson, D. & Ponta, H. *EMBO J.* **6**, 363-368 (1987).
- Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729-740 (1983).
- Gorman, C. M., Moffat, L. F. & Howard, B. H. *Mol. cell. Biol.* **2**, 1044-1051 (1982).

Essential role for phosphatidylinositol 4,5-bisphosphate in yeast cell proliferation

Isao Uno*, Kiyoko Fukami†, Hiroyuki Kato†, Tadaomi Takenawa† & Tatsuo Ishikawa*

*Institute of Applied Microbiology, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan
†Department of Pharmacology, Tokyo Metropolitan Institute of Gerontology, Itabashi-ku, Tokyo 173, Japan

The responses of mammalian cells to external signals are commonly mediated by intracellular secondary messengers, among which are the breakdown products of phosphatidylinositol 4,5-bisphosphate (PIP₂): 1,2-diacylglycerol (DG) and inositol 1,4,5-triphosphate (IP₃) (refs 1-7). Although phosphoinositide turnover in the yeast *Saccharomyces cerevisiae* has been shown to be regulated by glucose^{8,9} and sterol^{10,11}, as yet no definitive function has been ascribed to yeast phosphoinositides. We have recently developed a monoclonal antibody specific for PIP₂ and reported that it inhibits mitogenesis of mammalian cells stimulated by platelet-derived growth factor and bombesin¹². We now report that when introduced into yeast cells by electroporation this antibody inhibits their growth. Furthermore, several yeast mutants with temperature-dependent growth defects are altered in their sensitivity to our antibody and are found to have specific alterations in their phosphoinositide metabolism.

BALB/c mice were immunized with PIP₂ prepared from bovine spinal cords¹³, and a clone of hybridoma cells producing an antibody of immunoglobulin G_{2b} class was isolated. The immunoglobulin G preparation, designated antibody kt3g, had virtually no affinity for phosphatidylcholine, phosphatidylethanolamine, phosphatidylinositol (PI) or IP₃, and a weak but detectable affinity for phosphatidylinositol 4-phosphate (PIP), indicating specific binding to PIP₂ (ref. 12). To test the sensitivity of yeast cells to antibody kt3g, wild-type cells (IU87-1B) were suspended in YPGlu medium containing antibody kt3g, and were subjected to electroporation (electric field-mediated pro-

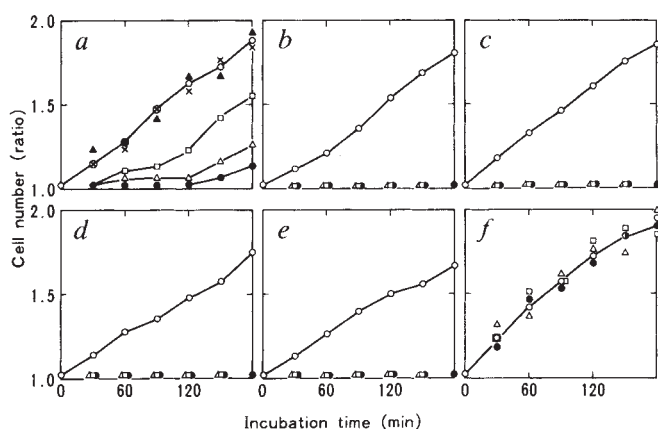


Fig. 1 Growth inhibition of wild-type and mutant cells by antibody kt3g. Cells of each strain were subjected to electroporation in the presence of 0 (○), 10 (□), 20 (△) and 50 (●) μ g of antibody kt3g. For the control, the wild-type cells were subjected to electroporation in the presence of non-immune mouse IgG_{2b} (50 μ g) (▲) or antibody kt3g (50 μ g) pretreated with PIP₂ (1 mM) (×). The treated cell suspension was incubated at 25 °C, and the cells were counted at the indicated times. At least 600 cells were counted for each value and to calculate the increase in cell number. *a*, IU87-1B (wild type); *b*, IU87-108 (*pim1*); *c*, IU87-103 (*pim2*); *d*, IU87-123 (*pim3*); *e*, IU87-158 (*pim4*); *f*, IU87-1 (*pim5*).

Methods. Cells of each strain (5×10^6 – 1×10^7 cells per ml at log phase) were suspended in YPD medium (final volume 1.0 ml) containing various amounts of antibody kt3g or non-immune mouse immunoglobulin IgG_{2b}. The cell suspensions were subjected to electroporation. Electroporation has been performed by using a Pro-Genator (Hoeffer Scientific Instrument); five square pulses (1 ms) at 700 V cm⁻¹ were given to the suspension. The introduction of IgG labelled with rhodamine or fluorescein isothiocyanate into cells by electroporation under the present conditions was confirmed by immunofluorescence microscopy.

tein transfer)^{14,15}. Growth of wild-type cells was inhibited significantly by electroporation in the presence of antibody kt3g (Fig. 1*a*). In the presence of 20 or 40 μ g of antibody kt3g, growth of wild-type cells was inhibited for 90 or 120 min, but then the growth rate recovered to almost normal, suggesting that antibody kt3g could be degraded by proteolytic enzymes after prolonged incubation. For the control, the wild-type cells were subjected to electroporation in the presence of non-immune mouse immunoglobulin G or antibody kt3g pretreated with PIP₂, but no inhibition of cell growth was observed (Fig. 1*a*). The results indicate that PIP₂ specifically bound with antibody kt3g in yeast cells could possibly be essential for the growth of the cells.

To select mutants sensitive to antibody kt3g, we first isolated 19 temperature-sensitive mutants which accumulate unbudded cells and fail to initiate DNA replication at 37 °C, and then tested the sensitivity of these mutants to antibody kt3g at 25 °C. Four out of 19 mutants (IU87-108, IU87-103, IU87-123, and IU87-158) obtained from the wild-type strain, SP1, were more sensitive to antibody kt3g than the wild-type strain at 25 °C (Fig. 1*b–e*). All mutants failed to grow for 3 h after electroporation in the presence of ≥ 20 μ g antibody kt3g. Each of these mutants was crossed to the wild-type strain and the resulting diploids were tested for their sensitivity to antibody kt3g. All diploids thus constructed showed sensitivity similar to that of the wild-type strain, indicating that these mutations are recessive to the wild-type counterparts. Complementation tests among the recessive mutants indicated that they all carry mutations at different loci, designated *pim1* (phosphoinositide metabolism), *pim2*, *pim3* and *pim4* respectively. During the mutant isolation we found a strain (IU87-1) that was more resistant to the antibody kt3g in our stock cultures (Fig. 1*f*). This mutation was recessive to the wild-type counterpart and the locus responsible for this mutation was designated *pim5*.

Table 1 Effect of substances associated with the phosphoinositide pathway on cell division in *pim* mutants

Drugs	Concentration (μ M)	Increase of cell number (%)			
		IU87-108 (<i>pim1</i>)	IU87-103 (<i>pim2</i>)	IU87-123 (<i>pim3</i>)	IU87-158 (<i>pim4</i>)
PIP ₂	100	97 ± 17	79 ± 15	9 ± 13	11 ± 9
PIP	100	11 ± 14	20 ± 15	8 ± 11	12 ± 8
PI	100	12 ± 11	14 ± 11	11 ± 7	9 ± 13
IP ₃	100	15 ± 13	12 ± 9	13 ± 6	8 ± 14
IP ₂	100	12 ± 10	13 ± 7	10 ± 7	15 ± 12
DOG	100	9 ± 14	11 ± 8	7 ± 9	13 ± 5
OAG	100	8 ± 11	12 ± 11	8 ± 13	11 ± 9
IP ₃ + DOG	100*	83 ± 19	75 ± 10	9 ± 12	9 ± 9
IP ₃ + OAG	100*	73 ± 17	75 ± 10	13 ± 11	9 ± 12
IP ₂ + DOG	100*	17 ± 9	16 ± 8	11 ± 9	11 ± 11
IP ₂ + OAG	100*	18 ± 10	17 ± 11	14 ± 7	10 ± 7

Cells of each mutant strain arrested at G1 state by the incubation at 37 °C for 3 h were subjected to electroporation in the presence of drugs as described in the legend of Fig. 1. These cultures were incubated at 37 °C for 3 h, and then the cell number was counted. At least 600 cells were counted to get each value and calculate the increase of cell number. The experiments were carried out five times, and the means were indicated with the standard errors.

Phosphatidylinositol 4,5-bisphosphate (PIP₂) was prepared from bovine spinal cords¹³, and inositol 1,4,5-triphosphate (IP₃) was prepared from PIP₂ by phospholipase C treatment. Phospholipase C that specifically degrades PIP₂ was obtained by purification from rat liver by the method of Nakanishi *et al.* (manuscript in preparation). Inositol 1,4-diphosphate (IP₂) was prepared as described by Irvine *et al.*¹⁷. Phosphatidylinositol (PI) and phosphatidylinositol 4-phosphate (PIP) were obtained from Sigma, and 1,2-dioctanoylglycerol (DOG) and 1-oleoyl-2-acetylgllycerol (OAG) from Nakarai Chem.

* The concentration of each reagent in the medium used for electroporation was 100 μ M.

To confirm that the mutants sensitive to antibody kt3g are defective in PIP₂ accumulation, PIP₂ was introduced into the mutant cells by electroporation. The *pim1* mutant cells incubated at the restrictive temperature for 3 h accumulated unbudded cells to more than 80%, but when the mutant cells received PIP₂ and electroporation, the number of unbudded cells decreased and the total number of cells increased about twofold (Fig. 2*a*). This indicates that the *pim1* mutant cells could divide once at the restrictive temperature in the presence of PIP₂. The effect of PIP₂ on the cell division of *pim1* mutant was dose-dependent, and 40 μ M PIP₂ was required for the half-maximum stimulation of cell division (Fig. 2*b*). The similar amount of PIP₂ was required for the stimulation of growth of *pim2* cells (Table 1). But cell division of *pim3* and *pim4* mutants was not affected by electroporation in the presence of PIP₂ (Table 1).

In mammalian cells both PIP and PIP₂ are produced from PI through sequential phosphorylation by the actions of PI kinase and PIP kinase, PIP₂ is degraded to produce DG and IP₃ by the action of phospholipase C, and IP₃ is dephosphorylated to inositol 1,4-diphosphate (IP₂) (ref. 16). To test whether these substances which are involved in phosphoinositide metabolism are effective in growth recovery of *pim* mutants, we introduced PI, PIP, IP₃ and synthetic DG, 1,2-dioctanoylglycerol (DOG) and 1-oleoyl-2-acetylgllycerol (OAG) singly or in combinations into the mutant cells by electroporation. None of these substances could stimulate cell division of the mutant cells, but the mixture of IP₃ and DOG or OAG could stimulate cell division of *pim1* and *pim2* mutant strains (Table 1). The mixture of IP₂ and DOG or OAG could not stimulate cell division of the mutants (Table 1). These results support an idea that PIP₂ may be degraded to produce IP₃ and DG for stimulation of yeast cell division.

To find the reaction affected by the mutations, we assayed PI kinase and PIP kinase in membrane and soluble fractions and the accumulation of PIP₂ (Table 2). About 80% of these enzyme

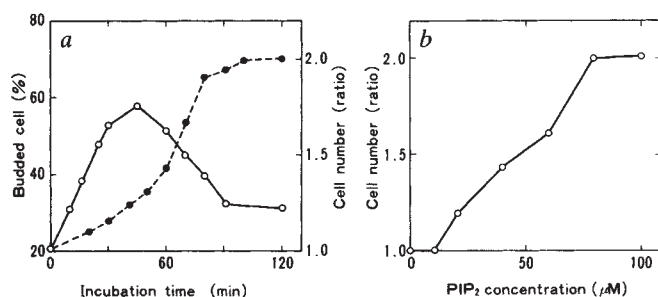


Fig. 2 Effect of PIP_2 on cell division of *pim1* mutant cells. *a*, Time course of increase of budded cells and total cell number. The *pim1* cells (5×10^6 cells per ml) at log phase were preincubated at 37°C for 3 h to arrest at G1 phase. These cells were subjected to electroporation in the presence of PIP_2 ($100 \mu\text{M}$) as described in Fig. 1 legend. These cells were incubated at 37°C and the number of budded cells (○) and the total number of cells (●) were counted at the indicated times. *b*, Effect of the PIP_2 concentration on cell division of *pim1* mutant cells. The *pim1* cells (5×10^6 cells per ml) at log phase were preincubated at 37°C for 3 h to arrest at G1 phase. These cells were subjected to electroporation in the presence of various amounts of PIP_2 and incubated at 37°C for 3 h, and the total number of cells was counted.

Table 2 Comparison of PI kinase and PIP kinase activities and accumulation of PIP_2 among wild-type and *pim* strains

Strain	Genotype	PI kinase activity (units mg^{-1} protein)		PIP kinase activity (units mg^{-1} protein)		PIP_2 (c.p.m./ 10^6 cells)	
		25°C	37°C	25°C	37°C	25°C	37°C
IU87-1B	+	4.2	5.5	85	103	923	1,018
IU87-108	<i>pim1</i>	4.1	5.5	85	44	1,041	442
IU87-103	<i>pim2</i>	3.3	1.2	85	100	1,243	324
IU87-123	<i>pim3</i>	1.9	2.2	40	46	764	549
IU87-158	<i>pim4</i>	1.4	1.3	24	30	812	571
IU87-1	<i>pim5</i>	10.9	13.2	136	144	2,411	2,340

The plasma membrane fraction was prepared from cells of each strain as described previously¹⁸. The PI kinase and PIP kinase activities of plasma membrane fractions were measured at 25°C and 37°C as described^{19,20}. Accumulation of PIP_2 was assayed by ^{32}P -incorporation into PIP_2 by the method of Talwalkar and Lester⁸. Cells ($\sim 1 \times 10^6$ cells per ml) were cultivated in the presence of $^{32}\text{P}_i$ (0.2 mCi ml^{-1}) at 25°C or 37°C for 30 min, and PIP_2 was extracted from the cells and purified by thin layer chromatography⁸. All values are means of values obtained from at least three independent experiments.

activities were found in membrane fractions and an essentially similar distribution of these enzymes in membrane and soluble fractions was observed in all mutant strains tested. The *pim1* mutant had normal PI kinase activity, but low PIP kinase activity at 37°C , and accumulated a low level of PIP_2 . The *pim2* mutant had low PI kinase activity and normal PIP kinase activity at 37°C , and accumulated a low level of PIP_2 . The *pim3* and *pim4* mutants had reduced activities of PI kinase and PIP kinase at both 25°C and 37°C , and produced lower levels of PIP_2 . But the *pim5* mutant produced increased activities of PI kinase and PIP kinase and accumulated significantly larger amounts of PIP_2 at both temperatures. These data suggest that *pim1* and *pim2* could be the genes for PI kinase and PIP kinase respectively, and that *pim3*, *pim4* and *pim5* may be related to the control of the PI turnover, but further biochemical and genetic studies are necessary to clarify the exact nature of these mutations.

This study using electroporation of yeast cells in the presence of anti- PIP_2 antibody or PIP_2 has presented evidence that a cellular event sensitive to anti- PIP_2 antibody, most likely PIP_2

breakdown to produce IP_3 and DG, is a crucial process in the mitotic cell cycle of yeast. The mutant study supported a conclusion that the production of PIP_2 and its breakdown may be essential to proceed the G1 stage of yeast cell cycle.

This work was supported in part by a Grant-in-Aid for Cancer Research from the Ministry of Education, Science and Culture of Japan.

Received 22 February; accepted 5 April 1988.

1. Michell, R. H. *Biochim. biophys. Acta* **415**, 81-147 (1975).
2. Takai, Y., Kishimoto, A., Kikkawa, U., Mori, T. & Nishizuka, Y. *Biochem. biophys. Res. Commun.* **91**, 1218-1224 (1979).
3. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
4. Nishizuka, Y. *Nature* **308**, 693-698 (1984).
5. Majerus, P. W., Neufeld, E. J. & Wilson, D. B. *Cell* **37**, 701-703 (1984).
6. Hokin, L. E. *Rev. Biochem.* **54**, 205-235 (1985).
7. Berridge, M. J. *Rev. Biochem.* **56**, 159-193 (1987).
8. Talwalkar, R. T. & Lester, R. L. *Biochim. biophys. Acta* **306**, 412-421 (1973).
9. Kaibuchi, K., Miyazaki, A., Arai, K. & Matsumoto, K. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8172-8176 (1986).
10. Dahl, J. S. & Dahl, C. E. *Biochim. biophys. Res. Commun.* **133**, 844-850 (1985).
11. Dahl, C., Biemann, H.-P. & Dahl, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4012-4016 (1987).
12. Matsuoka, K., Fukami, K., Nakanishi, O., Kawai, S. & Takenawa, T. *Science* **239**, 640-643 (1988).
13. Schacht, J. J. *Lipid Res.* **19**, 1063-1067 (1978).
14. Zimmermann, U., Pilwat, G. & Riemann, F. *Biochim. biophys. Acta* **375**, 209-219 (1975).
15. Wong, T.-K. & Neumann, E. *Biochem. biophys. Res. Commun.* **107**, 584-587 (1982).
16. Dawson, R. M. C. *Biochim. biophys. Acta* **33**, 68-77 (1959).
17. Irvine, R. F., Brown, K. D. & Berridge, M. J. *Biochem. J.* **222**, 269-272 (1984).
18. Matsumoto, K., Uno, I. & Ishikawa, T. *J. Bact.* **157**, 277-282 (1984).
19. Kai, M., White, G. L. & Hawthorne, J. N. *Biochem. J.* **101**, 328-337 (1965).
20. Kai, M., Salway, J. G. & Hawthorne, J. N. *Biochem. J.* **106**, 791-801 (1968).

Rates of primary electron transfer in photosynthetic reaction centres and their mechanistic implications

G. R. Fleming, J. L. Martin* & J. Breton†

Laboratoire d'Optique Appliquée, INSERM U275, Ecole Polytechnique-ENSTA, 91128 Palaiseau Cedex, France
† Service de Biophysique, CEN-Saclay, 91191 Gif-sur-Yvette Cedex, France

The conversion of light energy to chemical energy during photosynthesis involves the transfer of electrons between pigments embedded in a membrane protein. This process occurs with high quantum efficiency, the result of extremely fast electron transfer over a long distance preventing back transfer and energy loss. Recently the three-dimensional structures of the photosynthetic reaction centres of the bacteria *Rhodospseudomonas viridis*¹ and *Rhodobacter sphaeroides*² have been determined, allowing a molecular description of the primary charge separation process. There are two symmetrically related branches of pigments in the structure (L and M), extending from the special pair of bacteriochlorophyll molecules (P) to the two bacteriopheophytins (H_L and H_M) via two bacteriochlorophylls (B_L and B_M). Many features of the electron transfer process are poorly understood, such as the nature of the excited states involved, the identity of the primary charge separation step and the roles of the protein and of B^{3-13} . We have determined the rates of electron transfer in isolated reaction centre complexes of *Rps. viridis* and *Rb. sphaeroides* as a function of temperature. The rates increase as temperature is decreased, which may be due to either changes in electronic coupling of the pigments or changes in the population of coupled vibrational modes, or a combination of the two. We see no evidence of a B_L^- intermediate, which sets a lower limit on the rate of electron transfer from B_L to H_L . This is so high as to rule out transfer by two non-adiabatic steps.

Previous studies on isolated reaction centres have shown that the decay of the excited state of the special pair (P^*), monitored by stimulated emission in the infrared (around 1,045 nm for

* To whom correspondence should be addressed.

Rps. viridis and 925 nm for *Rb. sphaeroides*) provides an accurate measure of the electron transfer rate at room temperature¹⁴⁻¹⁶. We find that at all temperatures the stimulated emission signal from our samples closely follows a curve with a single exponential time constant. Figure 1 compares the stimulated emission decay for *Rps. viridis* with that of *Rb. sphaeroides* at 8 K. At this temperature the electron transfer time in *Rb. sphaeroides* is 1.2 ± 0.1 ps whereas in *Rps. viridis* it is 0.7 ± 0.1 ps.

Figure 2 compares the electron transfer rate over the temperature range 8–300 K and shows that this difference in electron transfer rates is temperature-dependent. At room temperature electron transfer times are similar in the two species. In both species there is an increase in rate as the temperature is lowered but the increase is much larger in *Rps. viridis*. Removal of B_M^{17} or complete deuteration of *Rb. sphaeroides* centres produce no detectable change in the electron transfer rate. We assume that the samples are in thermal equilibrium on the timescale of the electron transfer since no detectable change occurred in the decay of the stimulated emission upon excitation at 870 nm or at 580 nm.

The conventional expression for the rate of nonadiabatic electron transfer^{11,12} is:

$$k = (2\pi/\hbar) |V|^2 F \quad (1)$$

where V is the electronic interaction matrix element and F is the thermally averaged Franck-Condon factor. Both V and F may be temperature dependent, the former because protein contraction reduces the separation of donor and acceptor or changes the amount of charge transfer character in the initial state through relative motion of the special pair molecules⁴, the latter through changes in the population of vibrational modes that couple to the electron transfer. We first consider whether the temperature dependence of F can explain the increase in electron transfer with decreasing temperature. This approach has been used to discuss the temperature dependence of the second electron transfer step (H_L to Q_A)^{11,18} and the possible importance of coupling to polar modes of the protein has been suggested¹¹. A single mode representation for activationless electron transfer, for large displacement leads to¹¹:

$$k = k(0) \left[\frac{\exp(\hbar\omega/kT) - 1}{\exp(\hbar\omega/kT) + 1} \right]^{1/2} \quad (2)$$

where

$$k(0) = \frac{2\pi |V|^2}{\hbar^2 \omega (2\pi p)^{1/2}} \quad (3)$$

and ω is the average frequency of the coupled mode(s), $p = \Delta E/\hbar\omega$ and ΔE is the energy gap for the electron transfer. For this discussion we will assume ΔE is temperature independent for both species¹⁹. Equation (2) provides an excellent fit to the *Rb. sphaeroides* data with $k(0) = 8.33 \times 10^{11} \text{ s}^{-1}$ and $\hbar\omega = 80 \text{ cm}^{-1}$ (dashed line in Fig. 2). However the same expression gives a very poor description of the *Rps. viridis* data: to obtain a large enough change in rate a very low frequency (25 cm^{-1}) is required, but this produces a very poor fit (mixed line in Fig. 2) to the experimental data. Taking the fit of the *Rb. sphaeroides* data at face value and using⁷ $\Delta E = -1,290 \text{ cm}^{-1}$ we obtain a value of $V = 24 \text{ cm}^{-1}$.

The electronic coupling is also likely to be temperature dependent. Taking the data on myoglobin²⁰ as typical for thermal contraction of proteins the decreased distance between donor and acceptor should have little effect on the transfer rates, but if contraction increases the amount of charge transfer character⁴ in P^* then the interaction matrix element V should increase. Modelling of the optical spectrum indeed suggests that the interaction matrix element between the special pair molecules increases by about 100 cm^{-1} between 300 K and 10 K in both species³. The data in Fig. 2 could thus be interpreted according to equations (1) and (2) if V increased more strongly with decreasing temperature in *Rps. viridis* than in *Rb. sphaeroides*.

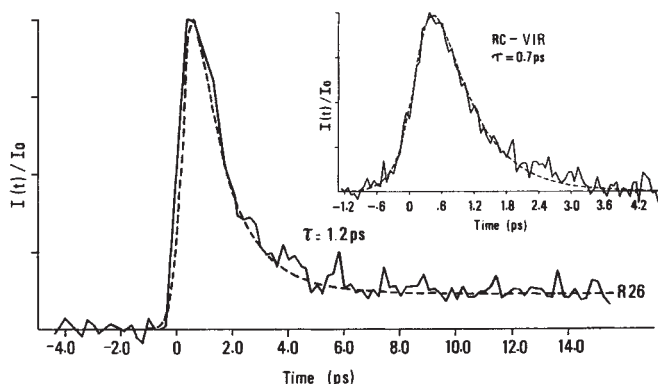
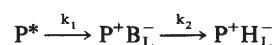


Fig. 1 Decay of stimulated emission for reaction centres of *Rb. sphaeroides* and *Rps. viridis* following excitation with 870 nm 150 fs pulses at 8 K. The experimental technique is described in refs 14 and 15. The samples were suspended in glycerol/buffer (60/40) and placed in a 1 mm cuvette. Temperature control was achieved using a helium gas flow convection cryostat (SMC, France). In *Rps. viridis* samples (RC-VIR) the probe wavelength is 1,045 nm and the dashed curve (---) corresponds to a single exponential fit with a time constant of 0.7 ps. In *Rb. sphaeroides* (R26), the probe wavelength is 915 nm. The best fit gives a time constant of 1.2 ± 0.1 ps.

This is consistent with EPR and ENDOR studies which indicate a less symmetrical electron distribution over the P triplet state in *Rps. viridis* than in *Rb. sphaeroides*²¹.

The involvement of the protein in the electron transfer process is of much current interest. Bixon and Jortner¹¹ conclude that for the later electron transfer step there is substantial coupling to low frequency polar modes of the protein. Theoretical models for hole burning experiments on reaction centres^{22,23} suggest strong coupling to modes with an average frequency of about 80 cm^{-1} (ref. 24). In view of the failure of equation (2) to describe the *Rps. viridis* data, it is not obvious how seriously to take the 80 cm^{-1} frequency obtained for all three *Rb. sphaeroides* samples. However, if expressions (1) and (2) do apply, even if V increases with decreasing temperature, this value of ω is a lower limit. Warshel²⁵ has estimated the intramolecular motion of the special pair to be $\approx 100 \text{ cm}^{-1}$. Higher frequency modes may also be coupled to the electron transfer. For $\hbar\omega \geq 200 \text{ cm}^{-1}$ the influence on the temperature dependence will be negligible; however, such modes will influence the absolute value of the rate. Our data on the fully deuterated sample rule out significant coupling to intramolecular C-H modes, protein X-H modes or protein-pigment hydrogen bonds in the case of *Rb. sphaeroides*.

The final result of the primary charge separation step is the formation of the state $P^+H_L^-$. The distance of closest approach between P and H_L is about $8 \text{ \AA}$ (refs 1, 2, 12). Conventional electron transfer theory in which V falls off with distance as $\exp[-\beta(r-r_0)]$, where r_0 is the distance at which the process becomes adiabatic (estimated as $3 \text{ \AA}$, ref. 12) and β has a value of $0.94\text{--}1.1 \text{ \AA}^{-1}$ (refs 26, 27) predicts rates that are much too slow. In *Rps. viridis* at 8 K, for example, the rate is only reduced by about 50% from the *adiabatic* rate expected if the relevant nuclear motion has a frequency of $\approx 100 \text{ cm}^{-1}$. This implies the existence of an almost perfectly conducting pathway to H_L , that is, a value for β of $\approx 0.05\text{--}0.1$ if the exponential form for V still holds. Marcus⁷ has proposed an alternative model in which an electron on B_L is a real intermediate state.



A variant from this model has also been proposed⁸ where a hole on B_L is a real intermediate ($B_L^+H_L^-$ state). In both models and for both steps $(r-r_0)$ is $1 \text{ \AA}$ and rates of the right order are predicted with conventional values of β . However, previous

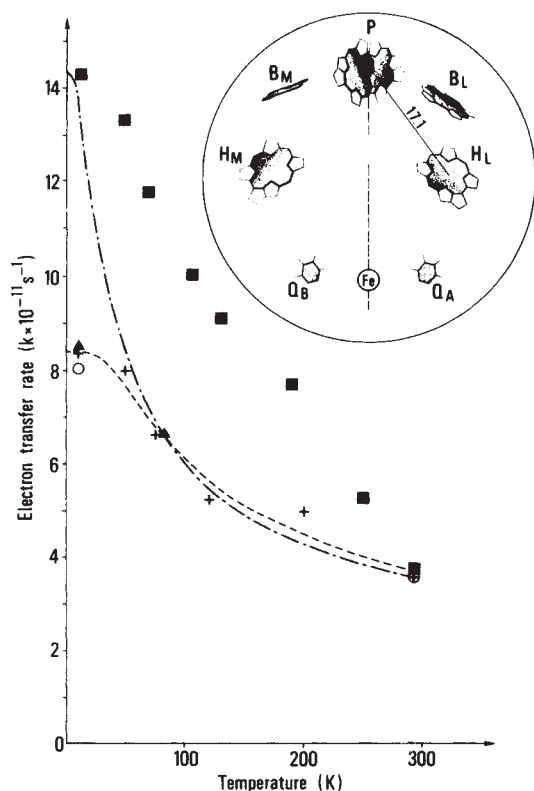


Fig. 2 Rate of primary electron transfer step as a function of temperature. ■, *Rps. viridis*. The mixed line (— · — · —) is a calculation using equation (2) with $\hbar\omega = 25 \text{ cm}^{-1}$ (see text). *Rb. sphaeroides*: +, unmodified; ○, modified (removal of B_M); ▲, deuterated samples. Dashed line, a calculation using equation (2) with $\hbar\omega = 80 \text{ cm}^{-1}$ which provides a good fit with the experimental data. The data for the modified and deuterated samples do not differ significantly from those for the unmodified sample (within our 10% precision limits). Inset: diagram of the structural relationships of the complex chromophores in the reaction centre^{1,2}. The centre-to-centre (P– H_L) distance is given in 10^{-1} nm . Q_A and Q_B , primary and secondary quinone electron acceptors, respectively.

experiments at room temperature did not reveal any bleaching in the absorption band of the B molecules^{14,15} and so both models postulate $k_2 > k_1$. Our low temperature data, where the spectrum is much better resolved, enable us to set new limits on the bleaching of B_L . Our analysis, described in detail elsewhere²⁸, leads to the estimate that k_2 must be 50–100 times k_1 for us to have failed to detect bleaching of B_L at 833 nm. This implies timescales of $\approx 10 \text{ fs}$ for the second step, clearly outside the regime of conventional electron transfer theory.

Thus any two-step process with both steps involving non-adiabatic electron transfer seems very unlikely. A single step mechanism such as one mediated by superexchange⁹ is not ruled out by our data. A third type of scheme in which non-adiabatic electron transfer occurs between P and B_L and adiabatic electron transfer occurs between strongly coupled B_L and H_L molecules has recently been suggested by Marcus²⁹. A possible test for this model would be to study the electron transfer rate in the presence of an applied electric field.

Part of this work has been supported by grants from INSERM and Ministère de la Recherche et de la Technologie. We thank Drs J. Norris and S. Kolaczowski for generously providing the deuterated reaction centres and J. W. Petrich for a critical reading of the manuscript. G.R.F. was a John Simon Guggenheim Fellow during the course of this work.

Received 14 January; accepted 28 March 1988.

- Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *J. molec. Biol.* **180**, 385–398 (1984).
- Allen, J. P., Fehér, G., Yeates, T. O., Komiyama, H. & Rees, D. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5730–5734 (1987).
- Scherer, P. O. J. & Fisher, S. F. *Biochim. biophys. Acta* **891**, 157–164 (1987).
- Parson, W. W. & Warshel, A. *J. Am. chem. Soc.* **109**, 6152–6163 (1987).
- Lous, E. J. & Hoff, A. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6147–6151 (1987).
- Won, Y. & Friesner, R. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5511–5514 (1987).
- Marcus, R. A. *Chem. Phys. Lett.* **133**, 471–477 (1987).
- Parson, W. W., Scherz, A. & Warshel, A. in *Antennas and Reaction Centers of Photosynthetic Bacteria* (ed. M. E. Michel-Beyerle) 122–126 (Springer, Berlin, 1985).
- Bixon, M., Jortner, J., Michel-Beyerle, M. E., Ogrodnik, A. & Lersch, W. *Chem. Phys. Lett.* **140**, 626–630 (1987).
- Chekalin, S. V., Matveev, Ya. A., Shkropotov, A. Ya., Shuvalov, V. A. & Yartsev, A. P. *FEBS Lett.* **216**, 245–248 (1987).
- Bixon, M. & Jortner, J. *J. phys. Chem.* **90**, 3795–3800 (1986).
- Marcus, R. A. & Sutin, N. *Biochim. biophys. Acta* **811**, 265–322 (1985).
- Friesner, R. & Wertheimer, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2138–2142 (1982).
- Martin, J. L., Breton, J., Hoff, A. J., Migus, A. & Antonetti, A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 957–961 (1986).
- Breton, J., Martin, J. L., Migus, A., Antonetti, A. & Orszag, A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5121–5125 (1986).
- Woodbury, N. W., Becker, M., Middendorf, D. & Parson, W. W. *Biochemistry* **24**, 7516–7521 (1985).
- Ditson, S. L., Davis, R. C. & Pearlstein, R. M. *Biochim. biophys. Acta* **766**, 623–629 (1984).
- Kirmaier, C., Holtz, D. & Parson, W. W. *Biochim. biophys. Acta* **810**, 33–48 (1985).
- Ogrodnik, A., Remy-Richler, N. & Michel-Beyerle, M. E. *Chem. Phys. Lett.* **135**, 576–581 (1987).
- Frauenfelder, H. et al. *Biochemistry* **26**, 254–261 (1987).
- Tiede, D. M. et al. in *The Photosynthetic Bacterial Reaction Center. Structure and Dynamics* (eds Breton, J. & Vermeglio, A.) NATO ASI Series (Plenum, New York, in the press).
- Boxer, S. G., Lockhart, D. & Middendorf, T. R. *Chem. Phys. Lett.* **123**, 476–482 (1986).
- Meech, S. R., Hoff, A. J. & Wiersma, D. A. *Chem. Phys. Lett.* **121**, 287–292 (1985).
- Hayes, J. M. & Small, G. J. *J. phys. Chem.* **90**, 4928–4931 (1986).
- Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3105–3109 (1980).
- Miller, J. R. in *Antennas and Reaction Centers of Photosynthetic Bacteria* (ed. Michel-Beyerle, M. E.) 234–241 (Springer, Berlin, 1985).
- Axup, A. W., Albin, M., Mayo, S. L., Crutchley, R. J. & Gray, H. B. *J. Am. chem. Soc.* **110**, 626–630 (1988).
- Martin, J. L., Breton, J., Lambry, J. C. & Fleming, G. R. in *The Photosynthetic Bacterial Reaction Center. Structure and Dynamics* (eds Breton, J. & Vermeglio, A.) NATO ASI Series (Plenum, New York, in the press).
- Marcus, R. A. *ibid.*

GUIDE TO AUTHORS

GENERAL

Please follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers for whom English is a second language and those who work in other fields. Please write clearly and simply, avoiding unnecessary technical terminology. *Nature's* staff will edit manuscripts to those ends if necessary. Contributors should check their proofs carefully. Because of the competition for space, many of the papers submitted for publication cannot be accepted. For this reason, and because brevity is a great assistance to readers, papers should be as brief as is consistent with intelligibility. Please note that one printed page of *Nature*, without diagrams or other interruptions of the text, has fewer than 1,300 words.

CATEGORIES OF PAPER

Review Articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance but must not exceed six pages of *Nature*.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles should be accompanied by a heading of 50–80 words written to advertise their contents in general terms, to which editors will pay particular attention. A heading (printed in italic type) is not an abstract and should not usually contain numbers or measurements. The study should be introduced in more detail in the first two or three paragraphs, which should also briefly summarize its results and implications.

Articles may contain a few subheadings of two or three words. The meaning of the text should not depend on the subheadings, whose function is to break up the text and to point to what follows. There should be fewer than 50 references.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should not have more than 1,000 words of text or more than four display items and should not occupy more than two pages of *Nature*. The first paragraph should describe, in not more than 150 words, the origins and chief conclusions of the study. Letters should not have subheadings or more than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned, most are unsolicited. They are normally between one and four pages of *Nature* in length.

News and Views articles are intended to inform non-specialist readers about a recently published advance. Suggestions should be made to the News and Views Editor. Illustrations are welcome. Proposals for meeting reports should be agreed in advance.

Scientific Correspondence is for the discussion of scientific matters, including contributions published in *Nature*. Priority is given to contributions of less than 500 words and five references. Figures are welcome.

Submission. All manuscripts may be submitted to either London or Washington. They should not be addressed to editors by name. Manuscripts by air courier sent to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax (VAT).

All contributions submitted for publication in *Nature* should conform with these rules.

Manuscripts should be typed, double-spaced, with a good-quality printer, on one side of the paper only. Four copies are required, each accompanied by lettered artwork. Four copies of half-tones should be included.

Reference lists, figure legends and so on should be on separate sheets, all of which should be double-spaced and numbered. Copies of relevant manuscripts in press or submitted for publication elsewhere should be included, clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their reference numbers.

Titles should say what the paper is about with the minimum of technical terminology and in fewer than 80 characters. Authors should avoid active verbs, numerical values, abbreviations and punctuation and should include one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name, figure number and, when known, manuscript reference number. Original artwork should be unlettered. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are permitted only if the parts are closely related, either experimentally or logically. Suggestions for cover illustrations, with captions, are welcome. Original artwork (and one copy of the manuscript) will be returned when a manuscript cannot be published.

Colour Artwork. A charge of £500 a page is made for 4-colour figures. Inability to meet these costs will not prevent the publication of essential colour figures if the circumstances are explained.

Figure legends must not exceed 300 words and ideally should be much shorter, and should use telegraphic form wherever possible. The figure should be described first, then, briefly, the method. Reference to a method described elsewhere is preferable to a full description. Methods should not be described in detail in the text.

References should be numbered sequentially as they appear in the text, followed by those in tables and finally those in the figure legends. Reference numbers apply only to papers published or in the press, and a different number should be given to each paper cited. All other forms of reference (including unrefereed abstracts) should be included in the text as personal communication, manuscript in preparation or preprint (with number and institution where appropriate). Text should not be included in the reference list. References should be abbreviated according to the *World List of Scientific Periodicals* (fourth edition, Butterworth, London, 1963–65). The first and last page numbers should be cited. Reference to books should clearly indicate the publisher and the date and place of publication.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided as much as possible and, when used, defined. Editors will shorten words if necessary. In case of doubt, SI units should be used.

Footnotes should not be used except for changed addresses or to identify the corresponding author if different from the first-named.

Acknowledgements should be brief. Grant numbers and contribution numbers are not required.